C 12 N 15/52 C 12 N 1/15 C 12 N 1/19 C 12 P 25/00

C 12 N 15/80

(19) BUNDESREPUBLIK DEUTSCHLAND

(i) DE 44 20 785 A 1



**DEUTSCHES** PATENTAMT Aktenzeichen:

P 44 20 785.9

Anmeldetag:

15. 6.94

Offenlegungstag:

5. 10. 95

② Erfinder:

Revuelta Doval, Jose Lui, Prof. Dr., Salamanca, ES; Santos Garcia, Maria Angeles, Dr., Santa Marta, ES; Buitrago Serna, Maria Jose, Salamanca, ES

(3) Innere Priorität: (2) (3) (3) 25.03.94 DE 44 10 382.4

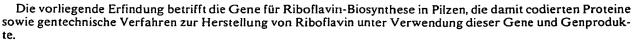
(71) Anmelder:

BASF AG, 67063 Ludwigshafen, DE

(54) Riboflavin-Biosynthese in Pilzen

Die vorliegende Erfindung betrifft die Gene für Riboflavin-Biosynthese in dem Pilz Ashbya gossypii sowie gentechnische Verfahren zur Herstellung von Riboflavin unter Verwendung dieser Gene und Genprodukte.

#### Beschreibung



Die Herstellung von Riboflavin durch Fermentation von Pilzen wie Eremothecium ashbyii oder Ashbya gossypii ist bekannt (The Merck Index, Windholz et al., eds. Merck & Co., Seite 1183, 1983).

In der EP 405370 sind Riboflavin-überproduzierende Bakterienstämme beschrieben, die durch Transformation der Riboflavin-Biosynthese-Gene aus Bacillus subtilis erhalten wurden.

Da die Genetik der Riboflavin-Biosynthese in Bakterien und Eukaryonten verschieden ist, sind die oben erwähnten Gene aus Bacillus subtilis nicht für ein rekombinantes Herstellverfahren für Riboflavin mit eukaryontischen Produktionsorganismen wie Ashbya gossypii geeignet.

In einer am 19.11. 1992 beim Deutschen Patentamt eingereichten Patentanmeldung wurde die Klonierung der Riboflavin-Biosynthese Gene der Hefe Saccharomyces cerevisiae beschrieben.

Eine Klonierung der Ashbya gossypii Riboflavin-Biosynthese Gene unter Verwendung der S. cerevisiae RIB-Gene mit üblichen Hybridisierungsmethoden gelang jedoch nicht; offenbar war die Homologie der RIB-Gene aus S. cerevisiae und A. gossypii für eine Hybridisierung nicht groß genug.

Es bestand daher die Aufgabe, die Riboflavin-Biosynthese Gene aus einem Eukaryonten zu isolieren, um damit ein rekombinantes Herstellverfahren für Riboflavin in einem eukaryontischen Produktionsorganismus bereitzustellen.

Demgemäß wurden in dem Ascomyceten Ashbya gossypii sechs Gene (rib-Gene), die für Enzyme der Riboflavin-Biosynthese ausgehend von GTP codieren, gefunden und isoliert.

Die Erfindung betrifft die folgenden DNA-Sequenzen:

DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 2 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO: 2, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäurensubstituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren.

DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 4 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO: 4, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren.

DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 6 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO: 6, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren.

DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 8 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO: 8, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren.

DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 10 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO: 10, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren.

DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 12 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO: 12, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren.

Die Gene und ihre Genprodukte (Polypeptide) sind im Sequenzprotokoll mit ihrer Primärstruktur aufgeführt und haben folgende Zuordnung:

SEQ ID NO:1: rib 1-Gen
SEQ ID NO:2: rib 1-Genprodukt (GTP-cyclohydrolase II)
SEQ ID NO:3: rib 2-Gen
SEQ ID NO:4: rib 2-Genprodukt (DRAP-Deaminase)
SEQ ID NO:5: rib 3-Gen
SEQ ID NO:6: rib 3-Genprodukt (DBP-Synthase)
SEQ ID NO:7: rib 4-Gen
SEQ ID NO:8: rib 4-Genprodukt (DMRL-Synthase)
SEQ ID NO:9: rib 5-Gen
SEQ ID NO:10: rib 5-Genprodukt (Riboflavin-Synthase)
SEQ ID NO:11: rib 7-Gen

SEQ ID NO: 12: rib 7-Genprodukt (HTP-Reductase)

Guanosintriphosphat (GTP) wird durch GTP-Cyclohydrolase II (rib 1-Genprodukt) zu 2,5-Diamino-6-ribosy-lamino-4-(3H)-pyrimidin-5-phosphat umgewandelt. Diese Verbindung wird anschließend durch rib 7-Genprodukt zu 2,5-Diamino-ribitylamino-2,4 (1H,3H)-pyrimidin-5-phosphat reduziert und dann durch rib 2-Genprodukt zum 5-Amino-6-ribitylamino-2,4 (1H,3H)-pyrimidindion deaminiert. Anschließend wird in einer rib 4-Genprodukt katalysierten Reaktion die C4-Verbindung DBP hinzugefügt und es entsteht 6,7-Dimethyl-8-ribityllumazin

10

15

20

35

40

(DMRL), aus dem in der rib 5-Genprodukt katalysierten Reaktion Riboflavin entsteht. Die C4-Verbindung DBP (L-3,4-Dihydroxy-2-butanon-4-phosphat) wird aus D-Ribulose-5-phosphat in einer rib 3-Genprodukt katalysierten Reaktion gebildet.

Die in SEQ ID NO: 1, 3, 5, 7, 9, 11 beschriebenen DNA-Sequenzen codieren für die Polypeptide, die in SEQ ID

NO: 2, 4, 6, 8, 10, 12 beschrieben sind.

Außer den im Sequenzprotokoll genannten DNA-Sequenzen sind auch solche geeignet, die infolge der Degeneration des genetischen Codes eine andere DNA Sequenz besitzen, jedoch für das gleiche Polypeptid codieren.

Weiterhin sind auch solche DNA Sequenzen Gegenstand der Erfindung, die für ein Genprodukt (Polypeptid) mit anderer als der im Sequenzprotokoll aufgeführten Primärstruktur codieren, solange das Genprodukt noch im wesentlichen die gleichen biologischen Eigenschaften wie das im Sequenzprotokoll genannte Genprodukt besitzt. Unter biologischen Eigenschaften sind vor allem die die Biosynthese von Riboflavin bewirkenden enzymatischen Aktivitäten zu verstehen.

Solche veränderten Genprodukte mit im wesentlichen gleichen biologischen Eigenschaften sind durch Deletion oder Hinzufügen von einer oder mehreren Aminosäuren oder Peptiden oder durch Austausch von Aminosäuren durch andere Aminosäuren erhältlich oder können aus anderen Organismen als Ashbya gossypii isoliert

Die DNA-Sequenzen, die für die veränderten Genprodukte codieren, sind zu den DNA-Sequenzen gemäß Sequenzprotokoll in der Regel zu 80 oder mehr Prozent homolog. Solche DNA-Sequenzen lassen sich ausgehend von den in SEQ ID NO: 1, 3, 5, 7, 9, 11 beschriebenen DNA-Sequenzen, beispielsweise mit üblichen Hybridisierverfahren oder der PCR-Technik aus anderen Eukaryonten als Ashbya gossypii isolieren. Diese DNA-Sequenzen hybridisieren unter Standardbedingungen mit den in SEQ ID NO: 1, 3, 5, 7, 9, 11 beschriebenen DNA-Sequenzen.

Unter Standardbedingungen sind beispielsweise Temperaturen zwischen 42 und 58°C in einer wäßrigen Pufferlösung mit einer Konzentration zwischen 0,1 und 1 × SSC (1 × SSC: 0,15 M NaCl, 15 mM Natriumcitrat pH 7,2) zu verstehen. Die experimentellen Bedingungen für DNA-Hybridisierungen sind in Lehrbüchern der Gentechnik, beispielsweise in Sambrook et al., "Molecular Cloning", Cold Spring Harbor Laboratory, 1989,

beschrieben.

Ein weiterer Gegenstand der Erfindung sind Regulationssequenzen, insbesondere Promotorsequenzen, die in 5-Richtung vor dem für das entsprechende Polypeptid codierenden DNA-Sequenzen liegen. Die Regulationssequenzen sind im Sequenzprotokoll aufgeführt und im folgenden näher erläutert.

Regulationssequenz für rib 1-Gen: SEQ ID NO: 1 Nukleotid 1-242 Regulationssequenz für rib 2-Gen: SEQ ID NO: 3 Nukleotid 1-450 Regulationssequenz für rib 3-Gen: SEQ ID NO: 5 Nukleotid 1-314 Regulationssequenz für rib 4-Gen: SEQ ID NO: 7 Nukleotid 1-270 Regulationssequenz für rib 5-Gen: SEQ ID NO: 9 Nukleotid 1-524 Regulationssequenz für rib 7-Gen: SEQ ID NO: 11 Nukleotid 1-352

Die Regulationssequenzen können auch noch in 5'- und/oder 3'-Richtung verkürzt werden, ohne daß ihre

Funktion wesentlich nachläßt. Essentiell für die Regulationswirkung sind in der Regel Fragmente von 30 bis 100, bevorzugt 40 bis 70

Nukleotiden aus den oben angegebenen Sequenzbereichen. Diese Regulationssequenzen können auch durch gerichtete Mutagenese im Vergleich zu den natürlichen Sequenzen in ihrer Funktion optimiert werden.

Die erfindungsgemäßen Regulationssequenzen eignen sich für die Überexpression von Genen in Ashbya, insbesondere von Genen, die für die Riboflavin-Biosynthese verantwortlich sind.

Weiterhin sind Gegenstand der Erfindung Expressionsvektoren, die eine oder mehrere der erfindungsgemä-Ben DNA-Sequenzen enthalten. Solche Expressionsvektoren erhält man, indem man die erfindungsgemäßen DNA-Sequenzen mit geeigneten funktionellen Regulationssignalen versieht. Solche Regulationssignale sind DNA-Sequenzen, die für die Expression verantwortlich sind, beispielsweise Promotoren, Operatoren, Enhancer, ribosomale Bindungsstellen, und die vom Wirtsorganismus erkannt und bedient werden.

Gegebenenfalls können noch weitere Regulationssignale, die beispielsweise Replikation oder Rekombination

der rekombinanten DNA im Wirtsorganismus steuern, Bestandteil des Expressionsvektors sein. Ebenso gehören die mit den erfindungsgemäßen DNA-Sequenzen oder Expressionsvektoren transformierten Wirtsorganismen zum Gegenstand der Erfindung. Bevorzugt werden als Wirtsorganismen eukaryontische Organismen, besonders bevorzugt solche der Gattung Saccharomyces, Candida, Pichia, Eremothecium oder Ashbya verwendet. Besonders bevorzugte Arten sind Saccharomyces cerevisiae, Candida flaveri, Candida

famata, Eremothecium ashbyii und Ashbya gossypii. Weiterhin gehört zur Erfindung ein rekombinantes Herstellverfahren für Riboflavin, in dem die erfindungsgemäßen transformierten Wirtsorganismen in an sich bekannter Weise durch Fermentation gezüchtet werden und das während der Fermentation gebildete Riboflavin aus dem Fermentationsmedium isoliert und gegebenenfalls

3

15

20

30

35

40

45

gereinigt wird.

5

10

30

50

Die rib-Gene und -Genprodukte lassen sich wie im Beispiel und im Sequenzprotokoll beschrieben isolieren und charakterisieren.

#### Beispiel 1

### Isolierung der Ashbya gossypii Riboflavin Biosynthese Gene (rib-Gene)

#### a. Konstruktion einer Ashbya gossypii cDNA-Bank

Gesamt RNA wurde aus dem Mycel des Riboflavin überproduzierenden Stammes Ashbya gossypii ATCC 10195 nach Züchtung auf YEPD Medium (Sherman et al., "Methods in yeast genetics", Cold Spring Harbor, New York, 1989) in der späten logarithmischen Wachstumsphase extrahiert.

Poly(A) † RNA wurde durch zweimalige Adsorption und Elution an oligo(dT)-Cellulose gereinigt (Aviv und Leder, Proc. Natl. Acad. Sci. USA 69, 1972, 1408—1412). Die cDNA wurde nach der allgemeinen Vorschrift von Gubler und Hoffmann isoliert (Gene 25, 1983, 263) und synthetische EcoRI-Adaptoren wurden an die Enden der bluntend cDNA-Moleküle hinzugefügt. Die EcoRI nachgeschnittenen cDNA Fragmente wurden anschließend mittels T4 Polynukleotidkinase phosphoryliert und in den dephosphorylierten EcoRI geschnittenen Vektor pYEura3 kloniert (Fig. 1). pYEura3 (Clonetech Laboratories, Inc., Kalifornien) ist ein Hefe-Expressionsvektor, der die Galaktose-induzierbaren GAL1 und GAL10 Promotoren und URA, CEN4 und ARS1 beinhaltet. Diese Hefeelemente erlauben die Transformation und Expression klonierter DNA-Fragmente in Hefezellen.

Aliquots der Ligationsreaktion wurden benutzt um hochkompetente (Hanahan, DNA Cloning, ed. D.M. Glover; IRL Press, Oxford 1985, 109) E. coli XL1-Blue (Bullock et al., Biotechniques 5 (1987) 376-378) zu transformieren und Transformanden wurden auf Basis ihrer Ampicillinresistenz selektioniert.

Etwa 3 × 10<sup>5</sup> ampicillinresistente Zellen wurden vereinigt, amplifiziert und daraus Plasmid-DNA isoliert (Birnboim und Doly, Nucleic Acids Res. 7, 1979, 1513).

### b. Isolierung von Ashbya gossypii cDNA-Klonen, die für riboflavinbildende Enzyme codieren

cDNA-Klone von Ashbya gossypii, die für riboflavinbildende Enzyme codieren, wurden durch funktionelle Komplementation von Saccharomyces cerevisiae Mutanten, die in der Riboflavin-Biosynthese betroffen sind, isoliert.

Die Stämme AJ88 (Mata leu2 his3 rib1::URA3 ura3-52), AJ115 (Matalpha leu2 inos1 rib2::URA3 ura3-52), AJ71 (Matalpha leu2 inos1 rib3::URA3 ura3-52), AJ106 (Matalpha leu2 inos1 rib4::URA3 ura3-52), AJ66 (Mata canR inos1 rib5::URA3 ura3-52) und AJ121 (Matalpha leu2 inos1 rib7::URA3 ura3-52) sind mutierte Stämme, die durch Zerstörung eines der sechs Gene (RIB1 bis RIB5 und RIB7), die in die Riboflavinbiosynthese bei Saccharomyces cerevisiae involviert sind.

Diese Stämme wurden jeweils mit 25 µg cDNA aus der Ashbya gossypii cDNA-Bank transformiert und auf festem Galaktose-haltigem Medium ohne Riboflavin ausplattiert. Nach ungefähr einer Woche Wachstum wurden Rib + Transformanden von den Kulturschalen isoliert.

Jeweils eine Transformande von jeder transformierten Mutante (Rib1+, Rib2+, Rib3+, Rib4+, Rib5+ und Rib7+) wurde analysiert und in allen Fällen wurde gefunden, daß der Rib+ Phänotyp nur in Galaktosemedium, nicht jedoch in Glucosemedium exprimiert war.

Diese Ergebnisse belegen, daß der Rib + Phänotyp unter der Kontrolle des plasmidständigen galaktoseinduzierbaren GAL10 Promotors exprimiert wurde.

Plasmid-DNA wurde aus den Rib1+, Rib2+, Rib3+, Rib4+, Rib5+ und Rib7+ Transformanden durch Transformation von E. coli isoliert und pJR715, pJR669, pJR788, pJR733, pJR681 und pJR827 genannt.

Partialsequenzierung der in diesen Plasmiden enthaltenen cDNA-Insertionen bestätigte, daß sie für Proteine codieren, die analog zu Proteinen der Rib-Genprodukte aus Saccharomyces sind.

### c. Isolierung von Ashbya gossypii genomischen Klonen, die für riboflavinbildende Enzyme codieren

Um die genomischen Kopien der riboflavinbildenden Gene von Ashbya gossypii zu isolieren wurde eine genomische Bank von Ashbya gossypii ATCC 10195 in dem Cosmid superCos1 (Stratagene Cloning Systems, Kalifornien) angelegt und mit <sup>32</sup>P-markierten Proben, die von den cDNA Kopien der RIB1, RIB2, RIB3, RIB4, RIB5 und RIB7 Gene von Ashbya gossypii abgeleitet waren, gescreent.

Cosmid Klone mit RIB1, RIB2, RIB3, RIB4, RIB5 und RIB7 DNA wurden isoliert durch Koloniehybridisierung (Grunstein und Hogness, Proc. Natl. Acad. Sci. USA 72, 1975, 3961—3965). Weitere Southern Analysen von enzymatisch gespaltener Cosmid DNA unter Verwendung der gleichen RIB-spezifischen cDNA Proben erlaubte die Identifizierung definierter Restriktionsfragmente, die die RIB1, RIB2, RIB3, RIB4, RIB5 und RIB7 Gene von Ashbya gossypii enthielten.

Ein 3,1 kb langes BamHI-ClaI DNA Fragment wurde gefunden, das das gesamte RIB1 Gen von Ashbya gossypii, codierend für GTP-Cyclohydrolase II enthielt. Dieses Fragment wurde aus einem Agarose Gel isoliert und in den BamHI und ClaI geschnittenen pBluescript KS (+) phagemid (Stratagene Cloning Systems) kloniert und lieferte so das Plasmid pJR765 (Fig. 2).

Eine 1329 bp lange DNA Sequenz wurde erhalten (SEQ ID NO: 1), die den RIB1 offenen Leserahmen von 906 bp, 242 bp von der 5'-nichtkodierenden Region und 181 bp von der 3'-nichtkodierenden Region enthielt.

Das gesamte Ashbya gossypii RIB2 Gen, das für die DRAP-Deaminase codiert, wurde auf einem 3,0 kb langen

DE 44 20 785

EcoRI-PstI Fragment gefunden, das kloniert in pBluescript KS (+) das Plasmid Pyk 758 ergab (Fig. 3). Eine 2627 bp lange Region der EcoRI-PstI-Insertion mit dem offenen Leserahmen von RIB2 von 1830 bp, 450 bp der 5'-untranslatierten Region und 347 bp der 3'-untranslatierten Region wurde sequenziert (SEQ ID NO:3). Das gesamte Ashbya gossypii RIB3 Gen, das für die DBP-Synthase codiert, wurde auf einem 1,5 kb langen Pstl-HindIII Fragment gefunden, das kloniert in pBluescript KS (+) das Plasmid PJR790 ergab (Fig. 4). Eine 1082 bp lange Region der PstI-HindIII-Insertion mit dem offenen Leserahmen von RIB3 von 639 bp, 314 bp der 5'-untranslatierten Region und 129 bp der 3'-untranslatierten Region wurde sequenziert (SEQ ID NO: 5). Das Ashbya gossypii RIB4 Gen, das für die DMRL-Synthase codiert, wurde auf einem 3,2 kb langen PstI-PstI Fragment gefunden, das kloniert in pBluescript KS (+) das Plasmid PJR762 ergab (Fig. 5).

Eine 996 bp lange Region der PstI-PstI-Insertion mit dem offenen Leserahmen von RIB4 von 519 bp, 270 bp der 5'-untranslatierten Region und 207 bp der 3'-untranslatierten Region wurde sequenziert (SEQ ID NO:7). Das gesamte Ashbya gossypii RIB5 Gen, das für die Riboflavin-Synthase codiert, wurde auf einem 2,5 kb langen PstI-PstI Fragment gefunden, das kloniert in pBluescript KS(+) das Plasmid PJR739 (Fig. 6) ergab.

Eine 1511 bp lange Region der PstI-PstI-Insertion mit dem offenen Leserahmen von RIB5 von 708 bp, 524 bp der 5'-untranslatierten Region und 279 bp der 3'-untranslatierten Region wurde sequenziert (SEQ ID NO:9). Schließlich wurde das Ashbya gossypii RIB7 Gen, das für die HTP-Reduktase codiert, auf einem 4,1 kb langen EcoRI-EcoRI-Fragment gefunden, das kloniert in pBluescript KS (+) das Plasmid PJR845 ergab (Fig. 7).

15

20

25

30

35

40

50

55

60

Eine 1596 bp lange Region der EcoRI-EcoRI-Insertion mit dem offenen Leserahmen von RIB7 von 741 bp, 352 bp der 5'-untranslatierten Region und 503 bp der 3'-untranslatierten Region wurde sequenziert (SEQ ID NO: 11).

#### Beispiel 2

### mRNA Analyse der Ashbya gossypii RIB-Gene

Um die RIB spezifischen Transkripte zu identifizieren wurden Northern Analysen durchgeführt. Gesamt RNA wurde aus dem Ashbya gossypii Stamm ATCC 10195 wie in Beispiel 1 beschrieben, isoliert. Die RNA Proben des Stammes (5 µg) wurden elektrophoretisch aufgetrennt auf 0,8% Agarose-Formaldehyd-Gelen zusammen mit RNA-Größenmarkern und unter Vakuum auf Nylonmembrane geblottet (Thomas, Proc. Natl. Acad. Sci. USA, 77, 1980, 5201 — 5205).

Die Nylonmembranen wurden unabhängig voneinander mit <sup>32</sup>P-markierten RIB-spezifischen DNA-Proben bei 42°C in 5 x SSC und in Gegenwart von 50% Formamid hybridisiert. Das Ashbya gossypii RIB1 Gen wird als unique Message von etwa 1150 Nukleotiden exprimiert, was in beiden Stämmen durch eine 0,7 kbp lange Smal-Sacl Probe aus dem Plasmid pJR765 (Fig. 8) nachgewiesen wurde.

Analog wurden unique 1900 Nukleotide lange RIB2-, 900 Nukleotide lange RIB3-, 800 Nukleotide lange RIB4-, 1050 Nukleotide lange RIB5- und 1000 Nukleotide lange RIB7-Transkripte in den Blots mit Hilfe eines 0,5 kbp langen Smal-Smal-Fragments aus pJR758, eines 0, 6 kbp langen HindIII-Kpnl-Fragments aus pJR790, eines 0,5 kbp langen Scal-HindIII Fragments aus pJR739 und eines 0,3 kbp langen PstI-PstI-Fragments aus pJR845 als spezifischer Probe nachgewiesen.

#### Beispiel 3

### Expression der Ashbya gossypii RIB-Gene in Saccharomyces cerevisiae

Wie in Beispiel 1 beschrieben, können gut untersuchte Mutanten von Saccharomyces cerevisiae, die in einer Stufe der Riboflavinbiosynthese defekt sind, auf Kulturmedien ohne Riboflavin wachsen, wenn sie ein Plasmid tragen, das für die komplementierenden Enzyme von Ashbya codiert. Um die Funktion der Ashbya gossypii RIB Genprodukte zu testen wurden flavinbildende Enzymaktivitäten in zellfreien Extrakten von S. cerevisiae-Mutanten gemessen, die eines der Expressionsplasmide pJR715, pJR669, pJR788, pJR733, pJR681 und pJR827

Diese in Beispiel 1 beschriebenen von pYEura3 abgeleiteten Plasmide enthalten Ashbya gossypii RIB-spezifische cDNA-Fragmente unter der Kontrolle des galaktoseinduzierbaren GAL10 Promotors.

Zellfreie Proteinextrakte von S. cerevisiae wurden aus Kulturen gewonnen, die in Flüssigmedium bis zu einer

optischen Dichte von etwa 2 OD gewachsen waren.

Die Zellen wurden geerntet, mit kaltem 20 mM Tris HCl, pH 7,5 gewaschen und im gleichen Puffer, der mit 1 mM Phenylethylsulfonylfluorid supplementiert war, resuspendiert.

Zell-Lysate wurden durch Vortexen in Gegenwart von Glaskugeln und Zentrifugation bei 3000 g für 20 min.

GTP-Cyclohydrolase II, DRAP-Deaminase, DBP-Synthase, DMRL-Synthase, Riboflavin-Synthase und HTPbei 4°C hergestellt. Reduktase Enzymaktivitäten wurden bestimmt wie in der Literatur beschrieben (Shavlovsky et al, Arch. Microbiol. 124 1980, 255-259; Richter et al., J. Baceriol. 175, 1993, 4045-4051; Klein und Bacher, Z. Naturforsch. 35b, 1980, 482-484; Richter et al. J. Bacteriol. 174, 1992, 4050-4056; Nielsen et al. J. Biol. Chem. 261, 1986, 3661; Plaut und Harvey, Methods Enzymol. 18B, 1971, 515 - 538; Hollander und Brown, Biochem. Biophys. Res. Commun. 89, 1979, 759 - 763; Shavlovski et al., Biochim. Biophys. Acta, 428, 1976, 611 - 618).

Protein wurde nach der Methode von Peterson quantifiziert (Anal. Biochem. 83, 1977, 346-356). Wie aus Tab. 1 ersichtlich, bewirkt das Plasmid pJR715 die Expression von GTP-Cyclohydrolase II Aktivität in der S. cerevisiae Mutante AJ88. Weiterhin ist diese Aktivität nur vorhanden in Zellen, die auf Galaktosemedium gewachsen sind, was darauf hinweist, daß die RIB1 cDNA Expression von Ashbya gossypii unter der Kontrolle des

galaktoseinduzierbaren GA romotors erfolgt.

Daher belegen diese Ergebnisse, daß RIB1 für die GTP-Cyclohydrolase II in Ashbya gossypii codiert. Auf analoge Art wurde gezeigt daß RIB2 für DRAP-Deaminase, RIB3 für DBP-Synthase, RIB4 für DMRL-Synthase, RIB5 für Riboflavinsynthase und RIB7 für HTP-Reduktase in diesem Pilz codiert.

Tabelle 1

GTP-Cyclohydrolase II Aktivität der S. cerevisiae RIB1 Mutante AJ88 und ihrer Transformanden

Stamm	Plasmid	GTP-Cyclohydrolase II U/mg Protein **)						
		Glucose	Galaktose					
X 2180-1A*	-	0,48	0,34					
AJ 88	-	n.d.	n.d.					
AJ 88	pIR715	n.d.	21,60					

n.d.:not detected

\*) Wildtyp

10

15

20

25

30

35

40

50

55

60

\*\*) Einheiten GTP-Cyclohydrolase II Aktivitäten
1U katalysiert die Bildung von 1 nmol HTP pro Stunde

Tabelle 2

DRAP-Deaminase Aktivität der S. cerevisiae RIB2 Mutante AJ115 und ihrer Transformanden

Stamm	Plasmid		eaminase otein *)
		Glucose	Galaktose
X 2180-1A		0,45	0,38
AJ 115	-	n.d.	n.d.
AJ 115	pIR669	n.d.	53,22

n.d.:not detected

\*) 1U katalysiert die Bildung von 1 nmol ARAP pro Stunde

Tabelle 3

DBP-Synthase Aktivität der S. cervisiae RIB3 Mutante AJ71 und ihrer Transformanden

Stamm	Plasmid		ynthase otein *)
		Glucose	Galaktose
X 2180-1A	_	0,80	0,75
AJ 71	-	n.d.	n.d.
AJ 71	pIR788	n.d.	25,19

n.d.:not detected

\*) 1U katalysiert die Bildung von 1 nmol DBP pro Stunde







Stamm	Plasmid	DMRL-S U/mg Pr	ynthase otein *)
		Glucose	Galaktose
X 2180-1A	-	2,04	1,73
AJ 106		n.d.	n.d.
AJ 106	pIR733	n.d.	86,54

n.d.:not detected

1U katalysiert die Bildung von 1 nmol DMRL pro Stunde \*)

Tabelle 5 Riboflavin-Synthase Aktivität der S. cerevisiae RIB5 Mutante AJ66 und ihrer Transformande

Plasmid	Riboflavi U/mg Pr	n-Synthase otein *)
	Glucose	Galaktose
	4,41	3,80
	n.d.	n.d.
DTP681	n.d.	164,20
	Plasmid pIR681	U/mg Pr Glucose - 4,41 - n.d.

n.d.:not detected

1U katalysiert die Bildung von 1 nmol Riboflavin pro Stunde \*)

Tabelle 6 HTP-Reduktase Aktivität der S. cerevisiae RIB7 Mutante AJ121 und ihrer Transformande

Stamm	Plasmid	HTP-Re U/mg Pr	ductase otein *)
		Glucose	Galaktose
. 0700 13	<del>_</del>	1,86	2,54
X 2180-1A	-	n.d.	n.d.
AJ 121 AJ 121	pIR827	n.d.	46,21

n.d.:not detected

RNSDOCID: DE 4420785A1:

1U katalysiert die Bildung von 1 nmol DRAP pro Stunde

60

5

10

15

20

25

30

35

40

45

50

### SEQUENZPROTOKOLL

	(1) ALGEMEINE INFORMATION:	
5	(i) ANMELDER:	
	(A) NAME: BASF Aktiengesellschaft	
	(B) STRASSE: Carl-Bosch-Strasse 38	
	(C) ORT: Ludwigshafen	
10	(E) LAND: Bundesrepublik Deutschland	
	(F) POSTLEITZAHL: D-67056	
	(G) TELEPHON: 0621/6048526	
	(H) TELEFAX: 0621/6043123	
15	(I) TELEX: 1762175170	
	(ii) ANMELDETITEL: Riboflavin-Biosynthese in Pilzen	
	(iii) ANZAHL DER SEQUENZEN: 12	
20	(iv) COMPUTER-LESBARE FORM:	
	(A) DATENTRÄGER: Floppy disk	
	(B) COMPUTER: IBM PC compatible	
25	(C) BETRIEBSSYSTEM: PC-DOS/MS-DOS	
23	(D) SOFTWARE: Patentin Release #1 0 Vorgion #1 25 (Trail	
	(2) INFORMATION ZU SEQ ID NO: 1:	
	(i) SEQUENZ CHARAKTERISTIKA:	
30	(A) LÄNGE: 1329 Basenpaare	
	(B) ART: Nukleinsäure	
	(C) STRANGFORM: Doppel	
	(D) TOPOLOGIE: linear	
35	(ii) ART DES MOLEKŪLS: CDNS zu mRNS	
	(iii) HYPOTHETISCH: NEIN	
	(iii) ANTISENSE: NEIN	
	(vi) URSPRŪNLICHE HERKUNFT:	
40	(A) ORGANISMUS: Ashbya gossypii	
	(ix) MERKMALE:	
	(A) NAME/SCHLÜSSEL: 5'UTR	
	(B) LAGE: 1242	
45	(ix) MERKMALE:	
	(A) NAME/SCHLÛSSEL: CDS	
	(B) LAGE: 2431148	
••	(ix) MERKMALE:	
50	(A) NAME/SCHLÜSSEL: 3'UTR	
	(B) LAGE: 11491329	
	(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:	
5	TTTCTGTCCG CATACTTCAT ATGCTCATCG CACATTGATA ATGTACATTC GAAAAATTTC	60
	TOUR TOUR TOUR THE TO	120
	AGGTGCTTTG TCATTCAACA ATCCACGTCG GAATTGGCGA CTATATAGTG TAGGGCCCAT	180
	AAAGCAGTAG TCGGTGTTGA TAGCTGTGTC AGACCAACTC TTTGTTAATT ACTGAAGCTG	240
0	AT ATG ACT GAA TAC ACA GTG CCA GAA GTG AGG TGT GTC GCA CGC GCG  Met Thr Glu Tyr Thr Val Bro Clu Hal	287
	Met Thr Glu Tyr Thr Val Pro Glu Val Arg Cys Val Ala Arg Ala  1 5 10	
	10 15	

												•						
CGC	ATA	CCG	ACG	GTA	CAG	GGC	ACC	GAT	GTC	TTC	CTC	CAT	CTA	TAC	CAC		335	
Arq	Ile	Pro	Thr	Val	Gln	Gly	Thr	Asp	Val	Phe	Leu	His	Leu	Tyr	His			
_				20					25					30				
AAC	TCG	ATC	GAC	AGC	AAG	GAA	CAC	CTA	GCG	ATT	GTC	$\mathtt{TTC}$	GGC	GAG	AAC	;	383	5
Asn	Ser	Ile	Asp	Ser	Lys	Glu	His	Leu	Ala	Ile	Val	Phe	Gly	Glu	Asn			
			35		-			40					45					
ልጥል	CGC	TCG		AGT	CTG	TTC	CGG	TAC	CGG	AAA	GAC	GAC	ACG	CAG	CAG		431	
								Tyr										10
110	*** 3	50					55	-		_	_	60						
CCG	CCC		CTC	CGG	GGC	GCC		GTG	GGC	CAG	CTG	TAC	CCC	GGG	CGG		479	
31a	Ara	Met	Val	Ara	Glv	Ala	Tvr	Val	Glv	Gln	Leu	Tyr	Pro	Gly	Arg			
ATG	65	Mec	V C2.3.	****	0_3	70	-2-				75	_						15
200		CCA	CAC	ccc	CAT		ССТ	CAG	GGC	CTG	GAG	CTG	CGG	TTT	GAT		527	
ACC The	CAG	312	3.00	λla	Acn	Ara	Ara	Gln	Glv	Leu	Glu	Leu	Arg	Phe	Asp			
	GIU	ALG	ASD	Ald	85	AL 9	y	·		90					95			
80			C3 C	CTC		CTIC	GAG	CGG	ece.		ACG	TGG	ACC	AGG	GAG		575	20
GAG	ACA	GGG	CAG		GIG	U-1	Clu	Arg	Δ1a	Thr	Thr	TTD	Thr	Ara	Glu			
GIU	יויחד	GIY	GIII		vaı	val	GIG	Arg	105	1111				110	•			
				100	OMO.	C3 C	mcc	GAG		መልሮ	acc.	GGC	GAG		GCG		623	
CCG	ACA	CTG	G.T.G	- ا	CTG	CAC	700	Clu	7.61	TAC	wh~	Gly	Glu	Thr	Δla		•	25
Pro	Thr	Leu		Arg	ren	HIS	Ser	Glu	Cys	TAT	1111	GLY	125	1111	n_u			
			115					120	<b>63.6</b>	mmc.	03.0	03.0		CCT	3 3 C		671	
TGG	AGC	GCG	CGG	TGC	GAC	TGC	GGG	GAG	CAG	TIC	GAC	CAG	21-	031	Tire		<b>J</b>	
Trp	Ser		Arg	Cys	qzA	Cys		Glu	GIN	Pne	Asp		ALa	GIY	гÃР			30
		130					135					140		~~~	03.0		719	
CTG	ATG	GCT	GCG	GCG	ACA	GAG	GGC	GAG	GTG	GTT	GGC	GGT	GCG	عادات	CAC		113	
Leu	Met	Ala	Ala	Ala	Thr		Gly	Glu	Val	Val		GIĀ	Ala	GIA	HIS			
	145					150					155			~~.	<b>220</b>		767	35
								GAG									767	
Gly	Val	Ile	Val	Tyr		Arg	Gln	Glu	Gly		Gly	Ile	GIY	Leu				
160					165					170					175		215	
GAG	AAG	CTG	AAG	GCG	TAC	AAC	CTG	CAG	GAC	CTG	GGC	GCG	GAC	ACG	GTG		815	40
Glu	Lys	Leu	Lys	Ala	Tyr	Asn	Leu	Gln		Leu	Gly	Ala	Asp		Val			
				180					185					190				
CAG	GCG	AAC	GAG	CTG	CTC	AAC	CAC	CCT	GCG	GAC	GCG	CGC	GAC	TTC	TCG		863	
Gln	Ala	Asn	Glu	Leu	Leu	Asn	His	Pro	Ala	Asp	Ala	Arg		Phe	Ser			45
			195					200					205					
TTG	GGG	CGC	GCA	ATC	CTA	CTG	GAC	CTC	GGT	ATC	GAG	GAC	ATC	CGG	TTG		911	
Leu	Gly	Arg	Ala	Ile	Leu	Leu	Asp	Leu	Gly	Ile	Glu	Asp	Ile	Arg	Leu			
		210					215					220						50
CTC	ACG	AAT	AAC	CCC	GAC	AAG	GTG	CAG	CAG	GTG	CAC	TGT	CCG	CCG	GCG		959	
Leu	Thr	Asn	Asn	Pro	Asp	Lys	Val	Gln	Gln	Val	His	Cys	Pro	Pro	Ala			
	225					230					235							
CTA	CGC	TGC	ATC	GAG	CGG	GTG	CCC	ATG	GTG	CCG	CTT	TCA	TGG	ACT	CAG	1	.007	55
Leu	Arg	Cys	Ile	Glu	Arg	Val	Pro	Met	Val	Pro	Leu	Ser	Trp	Thr	Gln			
240	_				245					250					255			

65

					(										)		
	CC	CAC	A CA	G, GGC	GT	3 - CG	C TC	G CG	C GA	G CT	G GA	C GG	C TAC	C 1	, 3 CG(	GCC	1055
	Pr	o Th	r Gl	a Gly	/ Va.	l Ar	g Se	r Ar	g Gli	ı Le	ı Ası	p Gly	/ Tyr	r Lei	ı Ard	y Ala	
					260	)				26	5				270	3	
5	AA	G GTY	C GA	G CGC	ATC	GGG	G CAC	YEA C	G CT	G CA	G CG	G CC	CTY	GTO	G CTC	G CAC	1103
	Ly	s Vai	l Glu	ı Arg	, Met	: Gly	/ His	s Met	Lei	ı Glı	ı Arg	g Pro	Leu	ı Val	l Lei	1 His	
				275	5				280	)				285	5		
	ACC	G TCT	r GCC	GCG	GCC	GAC	CTC	ccc	CGC	GCC	C AAC	ACA	CAC	ATZ	- እ ጥል!	ATCTTTGC	1155
10	Thi	Ser	: Ala	a Ala	Ala	Glu	ı Lei	ı Pro	Arg	, Ala	a Ası	ı Thi	His	I I I	 -		1133
			290	)				295					300		-		
	TAT	TTA!	AAA	CTCT	ATAA	AC G	TATO	CCAC	A CO	GCGC	CCG	GGG	CTGC	CAC	ACGO	CTGCTCA	1215
	CGC	GCTC	CCG	AACA	GTTC	TA A	CAAC	TAAT	rc go	GCGC	CTC	CCA	GTGA	TCG	TGGC	GAGCAC	1275
15	CTT	GTC	TCC	ATCA	TCAC	AT A	TCCI	CGGC	T AC	AGTO	GTC	TTC	AAGA	AGCG	TGC	i and the second	1329
	(2)	INF	ORMA	TION	ZU	SEQ	ID N	0: 2	:						+00	•	1329
				SEQU:													
				A) L													
20				B) A													
				D) To													
		(ii		T DE													
				QUEN						ח או							
25	Met	Thr	Glu	Tyr	Thr	Val	Pro	Glu	Val	D NO	~ ~ ~	77-1	×1 -	3		Arg	
	1				5					10		AGT	Ald	Arg			
	Ile	Pro	Thr	Val			ጥኮ፦	λen	V=1			T7.	T	<b></b>	15	Asn	
				20		C-7		Lop	25	FILE	rea	. mis	reu			Asn	
30	Ser	Ile	Asp		Lvs	Glu	Wie	Lou		T1.	**- 7		<b>~</b> 3.	30		Ile	
			35		<b>_</b> ,	Gra	1113	40	MIG	116	vai	Pue		GIu	Asn	Ile	
	Ara	Ser			T.011	Pho	λ <b>~</b> ~		3	7		_	45				
	3	50		Ser	200	FILE	55	TYL	Arg	гĀг	ASĐ		Thr	Gln	Gln	Ala	
35	Ara			Ara	G1 <sub>32</sub>	λ1 a		17-1	C1	<b>01</b> -	<b>.</b>	60	_				
	65			Arg	GLY	70	TYL	val	GTĀ	GIN		ıyr	Pro	GIA	Arg		
		Ala	Asp	Δla	Acn	_	λ <b></b>	C1-	<b>03</b>	<b>-</b>	75	_	_			80	
			1.05	Ala	85	ALY	Ary	GIII	GIY		GIU	Leu	Arg	Phe		Glu	
40	Thr	Glv	Gln	T.eu		17- 1	<b>C</b> 1	3		90		_			95		
			0411	Leu 100	Val	val	GIU	Arg		Thr	Thr	Trp	Thr		Glu	Pro	
	Thr	J. 211	Wa 1		T 011	774 -	<b>a</b>	-	105	_				110			
		200	115	Arg	Ter	MIS	Ser		Cys	Tyr	Thr	Gly		Thr	Ala	Trp	
45	Sar	<b>λ</b> 1 σ		~	3	<b>~</b>	- 1	120					125				
	Ser	130	ALG	Cys	ASD	Cys		GIU	Gln	Phe	Asp		Ala	Gly	Lys	Leu	
	Mot		31-		<b></b>		135					140					
	145	WIG	Ala	Ala	unr	GIU	Gly	Glu	Val	Val	Gly	Gly	Ala	Gly	His	Gly	
50		<b>-1</b> -	**- 1	_	_	150	_				155					160	
	val	TIE	vaı	Tyr	Leu	Arg	Gln	Glu	Gly		Gly	Ile	Gly	Leu	${ t Gly}$	Glu	
	T	<b>.</b>	-		165					170					175		
	ьys	ьеи	Lys	Ala	Tyr	Asn	Leu	Gln	Asp	Leu	${\tt Gly}$	Ala	Asp	Thr	Val	Gln	
55				180					185					190			
	Ala	Asn	Glu	Leu	Leu	Asn	His	Pro	Ala	Asp	Ala	Arg	Asp	Phe	Ser	Leu	
			192					200					205				
	GTĀ	Arg	Ala	Ile	Leu	Leu	Asp	Leu	Gly	Ile	Glu	Asp	Ile	Arg	Leu	Leu	
50		210					215					220		_			

Thr Asn Asn Pro Asp Lys Val Gln Gln Val His Cys Pro Pro Ala Leu	
225 230 235 240	
Arg Cys Ile Glu Arg Val Pro Met Val Pro Leu Ser Trp Thr Gln Pro	
245 250 255	5
Thr Gln Gly Val Arg Ser Arg Glu Leu Asp Gly Tyr Leu Arg Ala Lys	
260 265 270	
Val Glu Arg Met Gly His Met Leu Gln Arg Pro Leu Val Leu His Thr	
005	10
2/3	
Ser Ala Ala Glu Leu Pro Arg Ala Asn Thr His Ile	
230	
(2) INFORMATION ZU SEQ ID NO: 3:	15
(i) SEQUENZ CHARAKTERISTIKA:	
(A) LĀNGE: 2627 Basenpaare	
(B) ART: Nukleinsäure	
(C) STRANGFORM: Doppel	
(D) TOPOLOGIE: linear	20
(ii) ART DES MOLEKÜLS: cDNS zu mRNS	
(iii) HYPOTHETISCH: NEIN	
(iii) ANTISENSE: NEIN	
(vi) URSPRÜNLICHE HERKUNFT:	25
(A) ORGANISMUS: Ashbya gossypii	
(ix) MERKMALE:	
(A) NAME/SCHLÜSSEL: 5'UTR	
(B) LAGE: 1450	30
(ix) MERKMALE:	
(A) NAME/SCHLÜSSEL: CDS	
(B) LAGE: 4512280	35
(ix) MERKMALE:	
(A) NAME/SCHLÜSSEL: 3'UTR	
(B) LAGE: 22812627	
(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:	40
CTGCAGGACA ATTTAAATTA CGATTACACG CGGCAGCCTT CTTGGTGGGT. GTGGTTGGT	
TACAAGAATG ACCCCAAGCG GGTAAGAGTT CATAGGTATG CCTCCTTTCT TIGHTCTTCT	
ATTTTGAATT ATACTGATCA CGAACCCGTA ACGCTCGATG TCAGCGTTTC ATGCCATACA 180	
CAATTTGTCC CAATGGCTAT GCAGAATATT TCCCCACAGA GCACCATGGA AATGTATGTG 240	45
GGAGACGTCA CAGATATACT ACTGATGTTG TTCTCCAGAG TATACTACGC CCCTACCATA 300	
TTCGATCTTG TGGTATTGAC GATATTCCTC TGTTTGGTTT TACTGGCACT ATTCCGTTTG 360	
ACGGTATAGC GCTATTCGTT CATAGTGACA CATGCGGCAC TAGCTATTCA GCGAATCCTT 420	
TATAAACTGC TACTTAACGT TCGTAACACC ATG CTC AAA GGC GTT CCT GGC CTT 474	50
Met Leu Lys Gly Val Pro Gly Leu	
1 5	
CTT TTT AAG GAG ACG CAA CGT CAT CTG AAA CCC AGG CTG GTT AGG ATT 522	
Leu Phe Lys Glu Thr Gln Arg His Leu Lys Pro Arg Leu Val Arg Ile	55
10 15 20 ATG GAA AAC ACA TCG CAG GAT GAG AGT CGC AAA AGA CAG GTC GCT TCG 570	
ATG GAA AAC ACA TCG CAG GAT GAG AGT CGC AZZ HOLL CTG CTG	
Met Glu Asn Thr Ser Gln Asp Glu Ser Arg Lys Arg Gln Val Ala Ser	
25 30 35 40	60

	AAC	TTG	AGC	. AGC	GAI	GCC	GAT	GAG	GGC	TCG	CCG	GCA	GTT	ACG	AGG	CCG	618
	Asn	Leu	Ser	Ser	Asp	Ala	Asp	Glu	Gly	Ser	Pro	Ala	Val	Thr	Arg	Pro	
					45					50					55		
5	GTT	AAA	ATC	ACC	AAA	CGC	CTC	AGG	AAG	AAG	AAC	CTC	GGG	ACA	GGC	GAG	666
	Val	Lys	Ile		Lys	Arg	Leu	Arg	Lys	Lys	Asn	Leu	Gly	Thr	Gly	Glu	
	_			60					65					70			
	CTA	CGG	GAC	AAA	GCA	GGA	TTC	AAG	TTG	AAG	GTG	CAA	GAC	GTG	AGC	AAA	714
10	Leu	Arg		Lys	Ala	Gly	Phe	Lys	Leu	Lys	Val	Gln	Asp	Val	Ser	Lys	
			75					80					85				
	AAC	CGT	CAC	AGA	CAG	GTC	GAT	CCG	GAA	TAC	GAA	GTC	GTG	GTA	GAT	GGC	762
	Asn	Arg	His	Arg	Gln	Val	Asp	Pro	Glu	Tyr	Glu	Val	Val	Val	Asp	Gly	
15		90					95					100					
		ATG															810
		Met	Arg	Lys	Ile		Pro	Tyr	Phe	Phe		Tyr	Lys	Thr	Phe	Cys	
	105	~~~				110					115					120	
20		GAG															858
	гÃг	Glu	Arg	.r.p		Asp	Arg	Lys	Leu		Asp	Val	Phe	Val	Asp	Glu	
	ന്നസ	000	C3.0	000	125					130					135		
25	Dho	CGG	GAC 3.co	3	GAT	AGG	CCT	TAC	TAC	GAG	AAA	GTC	ATC	GGT	TCG	GGT	906
25	FIIG	Arg	ASD		ASD	Arg	Pro	Tyr		Glu	Lys	Val	Ile		Ser	${ t Gly}$	
	CCT	CTC	CTTC	140	330				145					150			
	Glv	GTG Val	Lau	LOU	AAC	Cl	AAG	TCA	TCG	ACG	TTA	GAT	AGC	GTA	TTG	CGT	954
30	GLy	Val	155	пец	MSII	GIÀ	гÃ2		ser	unr	Leu	Asp		Val	Leu	Arg	
	ААТ	GGA		CTC	מיניים ב	TCC	CAC	160	CELC	020	000	~~~	165				
	Asn	GGA G1v	Agn	T.All	TIO	100	Uia	Clu	CIG	CAC	CGT	CAT	GAG	CCA	CCG	GTC	1002
		Gly 170	-1.Op		110	SET	175	GIU	Ter	nis	Arg		GIU	Pro	Pro	Val	•
35	TCC	TCT	AGG	CCG	ATT	AGG		CTC	ጥልሮ	CAA	CAM	180	CNC	3.000	000	ama	1050
	Ser	Ser	Ara	Pro	Ile	Arg	Thr	Val	TUT	Glu	Acn	GAI	ACD	TIO	CTG	GIG	1050
	185					190			-1-	GIG	195	voħ	ASD	TIE	rea		
	ATT	GAC	AAG	CCC	AGC		ATT	CCA	GCC	САТ		ACC	ccc	CGT	ጥልሮ	200	1000
40	Ile	Asp	Lys	Pro	Ser	Gly	Ile	Pro	Ala	His	Pro	Thr	Glv	Ara	Tase.	y.c.	1098
					205	_				210				••••	215	mg.	
	TTC	AAC	TCC	ATT	ACG	AAA	ATA	CTT	GAA	AAA	CAG	CTT	GGA	TAC		CUT	1146
	Phe	Asn	Ser	Ile	Thr	Lys	Ile	Leu	Glu	Lys	Gln	Leu	Glv	Tvr	Thr	Val	1110
45				220					225					230			
	CAT	CCA	TGT	AAC	CGA	CTG	GAC	CGC	CTA	ACC	AGT	GGC	СТА		TTC	ጥጥር	1194
	His	Pro	Cys	Asn	Arg	Leu	Asp	Arg	Leu	Thr	Ser	Gly	Leu	Met	Phe	Leu	1174
			235					240				-	245				
50	GCA	AAA	ACT	CCA	AAG	GGA	GCC	GAT	GAG	ATG	GGT	GAT		ATG	AAG	GCG	1242
	Ala	Lys	Thr	Pro	Lys	Gly	Ala	Asp	Glu	Met	Gly	Asp	Gln	Met	Lvs	Ala	
		250					255					260					
55	CGC	GAA	GTG .	AAG	AAA	GAA	TAT	GTT	GCC	CGG			GGG	GAA	ттт	CCT	1290
55	Arg	Glu	Val	Lys	Lys	Glu	Tyr	Val	Ala	Arg	Val	Val	Gly	Glu	Phe	Pro	
	265					270					275		•			280	

60

												I					
ATA	GGT	GAG	ATA	GTT	GTG	GAT	ATG	CCA	CTG	AAG	ACT	ATA	GAG	CCG	AAG	1338	
Ile	Gly	Glu	Ile	Val	Val	Asp	Met	Pro	Leu	Lys	Thr	Ile	Glu	Pro	Lys		
				285					290					295			
CTT	GCC	CTA	AAC	ATG	GTT	TGC	GAC	CCG	GAA	GAC	GAA	GCG	GGC	AAG	GGC	1386	5
Leu	Ala	Leu	Asn	Met	Val	Cys	Asp	Pro	Glu	Asp	Glu	Ala	Gly	Lys	Gly		
			300					305					310				
GCT	AAG	ACG	CAG	TTC	AAA	AGA	ATC	AGC	TAC	${\tt GAT}$	GGA	CAA	ACG	AGC	ATA	1434	
Ala	Lys	Thr	Gln	Phe	Lys	Arg	Ile	Ser	Tyr	Asp	Gly	Gln	Thr	Ser	Ile		10
		315					320					325					
GTC	AAG	TGC	CAA	CCG	TAC	ACG	GGC	CGG	ACG	CAT	CAG	ATC	CGT	GTT	CAC	1482	
Val	Lys	Cys	Gln	Pro	Tyr	Thr	Gly	Arg	Thr	His	Gln	Ile	Arg	Val	His		
	330					335					340						15
TTG	CAA	TAC	CTG	GGC	TTC	CCA	ATT	GCC	AAC	GAT	CCG	ATT	TAT	TCC	AAT	1530	
Leu	Gln	Tyr	Leu	Gly	Phe	Pro	Ile	Ala	Asn	Asp	Pro	Ile	Tyr	Ser	Asn		
345					350					355					360		
CCG	CAC	ATA	TGG	GGC	CCA	AGT	CTG	GGC	AAG	GAA	TGC	AAA	GCA	GAC	TAC	1578	20
Pro	His	Ile	Trp	Gly	Pro	Ser	Leu	Gly	Lys	Glu	Cys	Lys	Ala	Asp	Tyr		
				365					370					375			
AAG	GAG	GTC	ATC	CAA	AAA	CTA	AAC	GAA	ATT	GGT	AAG	ACT	AAA	TCT	GCG	1626	26
Lys	Glu	Val	Ile	Gln	Lys	Leu	Asn	Glu	Ile	Gly	Lys	Thr	Lys	Ser	Ala		25
-			380					385					390				
GAA	AGT	TGG	TAC	CAT	TCT	GAT	TCC	CAA	GGT	GAA	GTT	TTC	AAA	GGG	GAA	1674	
Glu	Ser	Trp	Tyr	His	Ser	Asp	Ser	Gln	Gly	Glu	Val	Phe	Lys	Gly	Glu		30
		395					400					405				4500	30
CAA	TGC	GAT	GAA	TGT	GGC	ACC	GAA	CTG	TAC	ACT	GAC	CCG	GGC	CCG	AAT	1722	
Gln	Cys	Asp	Glu	Cys	Gly	Thr	Glu	Leu	Tyr	Thr	Asp	Pro	Gly	Pro	ASN		
	410					415					420					1770	35
GAT	CTT	GAC	TTA	TGG	TTG	CAT	GCA	TAT	CGG	TAT	GAA	TCC	ACT	GAA	CTG	1770	-
Asp	Leu	Asp	Leu	Trp	Leu	His	Ala	Tyr	Arg		Glu	Ser	Thr	GIU	Leu		
425					430					435					440	1010	
GAT	GAG	AAC	GGT	GCT	AAA	AAG	CGG	AGT	TAC	TCT	ACT	GCG	TTT	CCT	GAG	1818	40
Asp	Glu	Asn	Gly	Ala	Lys	Lys	Arg	Ser		Ser	Thr	Ala	Pne	Pro	GIU		
				445					450					455	010	1866	
TGG	GCT	CTT	GAG	CAG	CAC	GGC	GAC	TTC	ATG	CGG	CTT	GCC	ATC	GAA	CAG	1866	
Trp	Ala	Leu	Glu	Gln	His	Gly	Asp	Phe	Met	Arg	Leu	Ala	. 116	GIU	Gln		45
			460					465					470		CITIC	1914	
GCT	AAG	AAA	TGC	CCA	CCC	GCG	AAG	ACA	. TCA	TTT	AGC	GTT	GGI	GCC	G <b>TG</b>	1314	
Ala	Lys	Lys	Cys	Pro	Pro	Ala			Ser	Phe	ser	· Val	Gly	ATA	Val		
		475					480					485				1063	50
TTA	GTT	AAT	GGG	ACC	GAG	ATT	TTG	GCC	ACT	GGI	TAC	TCA	CGG	GAG	CTG	1962	
Leu	Val	Asn	Gly	Thr	Glu	Ile	Leu	Ala	Thr	Gly	Tyr	Ser	Arg	GIU	Leu		
	490					495					500					2010	
GAA	GGC	AAC	ACG	CAC	GCI	GAA	CAA	. TGI	· GCA	CTI	CAA	AAA	TAT	TTT	GAA	2010	55
Glu	Gly	Asn	Thr	His	Ala	Glu	Glr	Cys	Ala			Lys	Tyr	Pne	Glu		
505					510	)				515	•				520		

60

													- 4				
	CAA	CAT	AAA	ACC	GA	AAG	GTT	, CCI	ATT	GGT	ACA	GTA	ATA	TAC	ACG	ACT	2058
	Gln	His	Lys	Thr	Asp	Lys	Val	Pro	Ile	Gly	Thr	Val	Ile	Tvr	Thr	Thr	
					525					530				-2 -	535		
5	ATG	GAG	CCT	TGT	TCI	CTC	CGT	CTC	AGT	GGT	ААТ	<b>AAA</b>	CCG	ጥርጥ	COLOR	GAG	2106
	Met	Glu	Pro	Cys	Ser	Leu	Ara	Leu	Ser	Glv	Asn	Tage	Pro	Cve	Val	Glu	2106
				540			3		545		,,,,,,,	. Lys	110			GIU	
	CGT	АТА	АТС	_		CAG	CCT	ית א						550		GTA	
10	Ara	Tle	Ile	Cve	Gln	Cln	Clu	yen	71.	WC1	GCT	GIT	777	GIT	GGC	GTA	2154
10	3		555	Cy 3	GIII	Gill	GIY			Thr	ALA	var		Val	Gly	Val	
	CUM	CNC				~~~		560		_			565				
	CII	CIL	CCA	GAC	AAC	11C	GTG	AAG	AAC	AAT	ACA	AGT	CGT	GCG	CTA	TTG	2202
	neu	GIU	Pro	Asp	Asn	Pne		Lys	Asn	Asn	Thr	Ser	Arg	Ala	Leu	Leu	
15		570					575					580					
	GAA	CAA	CAT	GGT	ATA	GAC	TAT	ATT	CTT	GTC	CCT	GGG	TTT	CAA	GAA	GAA	2250
	Glu	Gln	His	Gly	Ile	Asp	Tyr	Ile	Leu	Val	Pro	Gly	Phe	Gln	Glu	Glu	
	585					590					595					600	
20	TGT	ACT	GAA	GCC	GCA	TTG	AAG	GGT	CAT	TGA'	rttt	GCT (	GCGA	ATTG'	ΓA		2297
	Cys	Thr	Glu	Ala	Ala	Leu	Lys	Gly	His								225,
					605		•	•	<b>-</b>	610							
	GAT	GACT'	TAA Z	AATA	rcga(	GG C	<b>ርጥልጥ</b> ን	ል ል ጥጥ	C GT(		للجلائك	አጥአሳ	מתי גיוו	nam /	~m» m/	STTTAC	2255
25	ATG	ACTG'	TTT 2	AAGC	ימכאיי	מנה טוד	ልጥልጥ <sup>ራ</sup>	1047-444 	ר אמר	בייים בייים		CCM	72001	TWI (	CIMI	ACGGTA	2357
	ATA	AATT	AAT (	BAGG	2 ልርጥ	ייי ייי	23337		C 22/	31 GW	211G	CCA	ATA:	rer :	1.66.17	ACGGTA	2417
	ACGO	ጋ ውጥ ነው። 	2AC 2	י שחשי	י ענויים ז	AC CC		1 T C G	D MM	CAA	ICTT	ATA:	PACG.	LTT (	JATG	ATATAA	2477
	AGT	יב ברבה מומול ל	DAC F		22 2 W	RG C	TACC.	IGAT	r Tr	TGC14	SAAC	TGT	l'TGT'	CAT A	\GGT:	TTTAC	2537
30	y www		IAG .		TAAG	rr re	5'1"1"1 <i>'E</i>	A'I"I'G'	r cc	CAG'	rcgg	CCA	ATTGT	erc (	CGGA	TTATT	2597
50			CCA :														2627
	(2)		ORMAT														
			(i) S														
25				l) Lā					iurer	1							
35				3) AF													
			(I	) TC	POL	GIE:	lir	near									
		(ii)	ART	DES	MOI	EKÜI	S: I	rote	ein								
		(xi)	SEC	UEN2	BES	CHRE	BUNG	3: SI	EQ II	NO:	4:						
40	Met	Leu	Lys	Gly	Val	Pro	Gly	Leu	Leu	Phe	Ivs	Glu	Thr	Gln	A ra	uic	
	1				5		_			10	-, -		****	0111	15	1113	
	Leu	Lys	Pro	Ara	Leu	Va 1	Ara	Tle	Mot		λcn	ωb~	Co	<b>01</b> -		<b>61.</b>	
		_		20			3		25	GIU	พวน	1111	Set		Asp	GIU	
45	Ser	Arσ	Lvs		Gln	t/a 1	7 T ¬	C		7	<b>a</b>	<b>a</b>	_	30	_		
		9	Lys 35	ALG	GIII	Val	AIG		ASN	rea	ser	ser		Ala	Asp	Glu	
	Clv	C0~		31-	**- 1	-1	_	40					45				
	GIY	267	Pro	Ald	var	unr	Arg	Pro	Val	Lys	Ile	Thr	Lys	Arg	Leu	Arg	
50	•	50	_	_	_		55					60					
	rys	Lys	Asn	Leu	Gly	Thr	Gly	Glu	Leu	Arg	Asp	Lys	Ala	Gly	Phe	Lys	
	65					70					75					80	
	Leu	Lys	Val	Gln	Asp	Val	Ser	Lys	Asn	Arg	His	Arg	Gln	Val	Asp	Pro	
55					85					90		_			95		
33	Glu	Tyr	Glu	Val	Val	Val	Asp	Gly	Pro	Met	Arq	Lys	Ile	Lvs	Pro	ጥ ጉ	
				100			_	-	105	-	3			110	~	-3 -	
	Phe	Phe	Thr	Tyr	Lys	Thr	Phe	Cvs		Glu	Δτσ	محدي	A ~~	7-70	7 ~~	T	
			115					120	_1 3		4	2		rap	ντα	⊓Ã 2	
60								14 U					125				

	130					135					140		Arg			
Tyr	Glu	Lys	Val	Ile	Gly	Ser	Gly	Gly	Val	Leu	Leu	Asn	Gly	Lys	Ser	
145		_			150					155					160	5
Ser	Thr	Leu	ASD	Ser	Val	Leu	Arg	Asn	Gly	Asp	Leu	Ile	Ser	His	Glu	
502				165					170					175		
•	***	3	uic		Pro	Pro	Va 1	Ser	Ser	Ara	Pro	Ile	Arg	Thr	Val	
ren	MIS	ALG		GIU	FIU		141	185		3			190			10
	_		180	_			**- 3		200	T 7.6	Dro.	Car		Tla	Pro	
Tyr	Glu	Asp	Asp	ASD	TIE	rea		TIE	Asp	гÃг	PIO	205	Gly	110		
		195					200	_	_	_			<b>T</b>	<b>+1</b> -	T	
Ala	His	Pro	Thr	Gly	Arg	Tyr	Arg	Phe	Asn	ser	TTE	THE	Lys	TIE	Leu	
	210					215					220				_	15
Glu	Lys	Gln	Leu	Gly	Tyr	Thr	Val	His	Pro	Cys	Asn	Arg	Leu	Asp	Arg	
225					230					235					240	
Leu	Thr	Ser	Gly	Leu	Met	Phe	Leu	Ala	Lys	Thr	Pro	Lys	Gly	Ala	Asp	
				245					250					255		20
C1.,	Mot	G114	Δεπ		Met	Lvs	Ala	Arσ	Glu	Val	Lys	Lys	Glu	Tyr	Val	
GIU	Mec	GTY	260	<b>G</b> 2		-20		265			•	_	270			
		7	200	<b>~1</b>	<b>~</b> 1	Dho	D~o		Clar	Glu	710	V=1	Val	ASD	Met	
Ala	Arg		var	GIY	GIU	Pile		TTE	Gry	Gra		285				25
		275			_		280			•	<b>.</b>		17m 7	~	A cm	
Pro	Leu	Lys	Thr	Ile	Glu		Lys	Leu	Ala	ren	ASII	mer	Val	Cys	Yah	
	290					295				_	300		_	_	-1-	
Pro	Glu	Asp	Glu	Ala	Gly	Lys	Gly	Ala	Lys		Gln	Phe	Lys	Arg	Tie	30
305					310					315				_	320	30
Ser	Tyr	Asp	Gly	Gln	Thr	Ser	Ile	Val	Lys	Cys	Gln	Pro	Tyr	Thr	Gly	
				325					330					335		
Arg	Thr	His	Gln	Ile	Arg	Val	His	Leu	Gln	Tyr	Leu	Gly	Phe	Pro	Ile	25
_			340					345					350			35
Ala	Asn	asa	Pro	Ile	Tyr	Ser	Asn	Pro	His	Ile	Trp	Gly	Pro	Ser	Leu	
		355					360					365				
Glv	Tare		Cvs	LVS	Ala	Asp	Tvr	Lys	Glu	Val	Ile	Gln	Lys	Leu	Asn	•
013	370		-1 -	-1 -	•	375	•	•			380					40
<b>~1</b>		C111	Tue	<b>ም</b> ክ ም	Tare		Δla	Glu	Ser	Tro	Tvr	His	Ser	GZA	Ser	
	TIE	GIY	пуз	1111	390	501	****			395	-1-			-	400	
385			•	-1	-	<b>61</b>	<b>6</b> 1	<b>01</b> -	~		. (23)	Circ	Glaz	Thr		
Gln	Gly	Glu										Cys	Gly	415		45
												_				
Leu	Tyr	Thr	Asp	Pro	Gly	Pro	Asn			Asp	Leu	.r.rp	Leu	HIS	Ald	
			420					425					430			
Tyr	Arg	Tyr	Glu	Ser	Thr	Glu	Leu	Asp	Glu	Asn	Gly		Lys	Lys	Arg	50
		435					440					445				
Ser	Tyr	Ser	Thr	Ala	Phe	Pro	Glu	Trp	Ala	Leu	Glu	Gln	His	Gly	Asp	
	450					455					460					
Dhe		<b>Δ ~</b> α	T,@11	<b>ء 1</b> ۵	T1 =		Gln	Ala	Lvs	Lvs	Cvs	Pro	Pro	Ala	Lys	
		AL Y	neu.		470		~~~		-, -	475		. = -	_		480	55
465		<b>71</b>	<b></b>	77-7		<b>73</b> -	77-7	Len	₹ <b>7</b> = 7			ጥኮታ	Glii	Ile	Leu	
TOT	ser	rue	ser			vrq	AGI	rea			229			495		
				485					490					793		

					•											
		Thr Gly	200					505					E 1 A			
5		la Leu 515					520					525	Lys			
	ر	ly Thr 30				535					540	Ser				
10	242	ly Asn			550					555					E C O	
		hr Ala		565					570					575	Lys	
15		sn Thr	280					585					500			
		al Pro 595	Gly	Phe	Gln	Glu	Glu 600	Cys	Thr	Glu		Ala 605	Leu	Lys	Gly	
20	His															
	(2) II	NFORMAT (i) SEC	UENZ	CHA	RAKT	ERIS	TIKA	:								
25	(A) LÂNGE: 1082 Basenpaare  (B) ART: Nukleinsaure  (C) STRANGFORM: Doppel  (D) TOPOLOGIE: linear  (ii) ART DES MOLEKÜLS: cDNS zu mRNS  (iii) HYPOTHETISCH: NEIN															
30	(ii (ii	.i) ART .i) HYP .i) ANT ri) URS	OTHE ISEN	TISCE SE: 1	H: N NEIN	EIN		zu m	RNS							
35		(A X) MER (A	) OR KMAL ) NAI	GANIS	MUS CHLŪS	: Asi	hbya		sypi:	i						
40		x) MER (A (B	KMAL: ) NA! ) LA(	E: Æ/SC GE: 3	HLŪS	SSEL	: CDS	5								
45	(x:	B) i) SEQI	) NAM ) LAC JENZE	ME/SC SE: 9 SESCH	54	1082 SUNG:	SEC	O ID	NO:	5:						
50	CCCTTC: AATGTA( CAAGTG!	ITGC AC CCAC C: AAAC A:	CGGT( FATC) FATC	GTTT GTAG CATC	CTC TTT GCC	AAAC ACTA AGCA	TCT TCG GGT	ACGA GATT	TACT CTACT	TG C	TAAC	AGC?	NG AC	CTGT	TAGG	60 120 180
55	TCACCTA ATACGTA ACAAACO	ACAA AA	AATTI	CAAC	GTI CA A	TTAC GC C	:AAG :CA I	TTCC GC A	CAAG CT G	CT I	AGTO	AACI	ra or o oo	CACC	AACG	240 300 350
				1				5		_			10			

16

60

				•								•					
GAG	CAG	TTC.	AAG	CAA	AAT	AAG	ATG	ATC	ATC	GTC	ATG	GAC	CAC	ATC	TCG	398	
Glu	Gln	Phe	Lys	Gln	Asn	Lys	Met	Ile	Ile	Val	Met	Asp	His	Ile	Ser		
		15	•				20					25					
AGA	GAA	AAC	GAG	GCC	GAT	CTA	ATA	TGT	GCA	GCA	GCG	CAC	ATG	ACT	GCC	446	5
Arg	Glu	Asn	Glu	Ala	Asp	Leu	Ile	Cys	Ala	Ala	Ala	His	Met	Thr	Ala		
_	30					35					40						
GAG	CAA	ATG	GCA	TTT	ATG	ATT	CGG	TAT	TCC	TCG	GGC	TAC	GTT	TGC	GCT	494	
Glu	Gln	Met	Ala	Phe	Met	Ile	Arg	Tyr	Ser	Ser	Gly	Tyr	Val	Cys	Ala		10
45					50					55					60		
CCA	ATG	ACC	AAT	GCG	ATT	GCC	GAT	AAG	CTA	GAC	CTA	CCG	CTC	ATG	AAC	542	
Pro	Met	Thr	Asn	Ala	Ile	Ala	Asp	Lys	Leu	Asp	Leu	Pro	Leu	Met	Asn		
	٠.			65					70					75			15
ACA	TTG	AAA	TGC	AAG	GCT	TTC	TCC	GAT	GAC	AGA	CAC	AGC	ACT	GCG	TAT	590	
Thr	Leu	Lys	Cys	Lys	Ala	Phe	Ser	Asp	Asp	Arg	His	Ser	Thr	Ala	Tyr		
			80					85					90				20
ACA	ATC	ACC	TGT	GAC	TAT	GCG	CAC	GGG	ACG	ACG	ACA	GGT	ATC	TCC	GCA	638	20
Thr	Ile	Thr	Cys	Asp	Tyr	Ala	His	Gly	Thr	Thr	Thr	Gly	Ile	Ser	Ala		
		95					100					105					
CGT	GAC	CGG	GCG	TTG	ACC	GTG	AAT	CAG	TTG	GCG	AAC	CCG	GAG	TCC	AAG	686	25
Arg	Asp	Arg	Ala	Leu	Thr	Val	Asn	Gln	Leu	Ala	Asn	Pro	Glu	Ser	Lys		23
	110					115					120						
GCT	ACC	GAC	TTC	ACG	AAG	CCA	GGC	CAC	ATT	GTG	CCA	TTG	CGT	GCC	CGT	734	
Ala	Thr	Asp	Phe	Thr	Lys	Pro	Gly	His	Ile	Val	Pro	Leu	Arg	Ala	Arg		30
125					130					135					140		30
GAC	GGC	GGC	GTG	CTC	GAG	CGT	GAC	GGG	CAC	ACC	GAA	GCG	GCG	CTC	GAC	782	
Asp	Gly	Gly	Val	Leu	Glu	Arg	Asp	Gly	His	Thr	Glu	Ala	Ala	Leu	Asp	•	
				145					150					155			35
TTG	TGC	AGA	CTA	GCG	GGT	GTG	CCA	GAG	GTC	GCT	GCT	ATT	TGT	GAA	TTA	830	
Leu	Cys	Arg	Leu	Ala	Gly	Val	Pro	Glu	Val	Ala	Ala	Ile		Glu	Leu		
			160					165					170		. = .	079	
GTA	AGC	GAA	AGG	GAC	GTC	GGG	CTG	ATG	ATG	ACT	TTG	GAT	GAG	TGT	ATA	878	40
Val	Ser	Glu	Arg	Asp	Val	Gly		Met	Met	Thr	Leu		Glu	Cys	IIe		
		175					180					185	G 3 M	030	cmc	926	
GAA	TTC	AGC	AAG	AAG	CAC	GGT	CTT	GCC	CTC	ATC	ACC	GTG	CAT	GAC	CTG	920	
Glu	Phe	Ser	Lys	Lys	His	Gly	Leu	Ala	Leu	ire		vaı	HIS	ASD	ned		45
	190					195					200			s cmc	CCECE	980	
									ACGG	CAA	CGAG	11C1	TT A	AGIC	GGTGT	360	
Lys	Ala	Ala	Val	Ala		Lys	Gln										
205					210						- C- D-1		CM3	~ മന്ന	መእመሮአ እ	1040	50
												TGAA	CIA	GMII	TATCAA	1082	
						TTCG			.I.I.W.I	CGAI	AC					1002	
(2)						ID N											
						AKTE			_								5 <b>5</b>
						2 Am		aure	71								
		•				osāu											
	, , ,					: li											60
	(11	) AR	T DE	S MO	LEKU	LS:	STOC	GTII									90

		(xi	) SE	QUE	2BES	CHRE	IBUN	G: S	EQ I	ON G	: 6:					
	Met	Thr	Ser	Pro	Cys	Thr	Asp	Ile	Gly	Thr	Ala	Ile	Glu	Gln	Phe	Tvs
	1				5				_	10					15	
5	Gln	Asn	Lys	Met	Ile	Ile	Val	Met	Asp	His	Ile	Ser	Arq	Glu		Glu
				·20					25					30		
	Ala	Asp	Leu	Ile	Cys	Ala	Ala	Ala	His	Met	Thr	Ala	Glu	Gln	Met	Ala
			35					40					45			
10	Phe	Met	Ile	Arg	Tyr	Ser	Ser	Gly	Tyr	Val	Cys	Ala	Pro	Met	Thr	Asn
		50					55					60				
	Ala	Ile	Ala	Asp	Lys	Leu	Asp	Leu	Pro	Leu	Met	Asn	Thr	Leu	Lys	Cys
	<sub>.</sub> 65					70					75					80
15	Lys	Ala	Phe	Ser	Asp	Asp	Arg	His	Ser	Thr	Ala	Tyr	Thr	Ile	Thr	Cys
					85					90					95	
	Asp	Tyr	Ala	His	Gly	Thr	Thr	Thr	Gly	Ile	Ser	Ala	Arg	Asp	Arg	Ala
				100		•			105					110		
20	Leu	Thr	Val	Asn	Gln	Leu	Ala	Asn	Pro	Glu	Ser	Lys	Ala	Thr	Asp	Phe
			115					120					125			
	Thr	Lys	Pro	Gly	His	Ile	Val	Pro	Leu	Arg	Ala	Arg	Asp	Gly	Gly	Val
95	_	130	_		_		135					140				
25	Leu	Glu	Arg	Asp	Gly		Thr	Glu	Ala.	Ala	Leu	qzA	Leu	Cys	Arg	Leu
	145			_		150					155					.160
	Ala	Gly	Val	Pro		Val	Ala	Ala	Ile		Glu	Leu	Val	Ser	Glu	Arg
30	•			_	165					170					175	
30	Asp	Val	Gly		Met	Met	Thr	Leu		Glu	Cys	Ile	Glu	Phe	Ser	Lys
	T	•••		180		_			185					190		
	гĀ2	HIS	Gly	Leu	Ala	Leu	Ile		Val	His	Asp	Leu	Lys	Ala	Ala	Val
35	21.		195	<b>63</b> .				200					205			
	Ald	210	Lys	GIN												
	(2)		\D\$63.00	17031	G-17 G			_								
	(4)		RMAT													
40		(1)	SEQ													
					NGE: T: N				are							
					RANG											
					POLO				•							
45		(ii)	ART						~ ~	סזגם						
	(	;, iii)	HYP	OTHE	ייין פרנייי	H. N	ETNI	באום.	2U III	1442			i			
			ANT													
	•		URS					ET.								
50		,			GANI											
		(ix)	MER	KMAT	E.	SMUS	- AS	ımya	. gos	sypi	1					
		( ==== ,			ME/S	Cut.ñ	CCTT	. 5/	amti							
					ME/S .GE:			. 5	UIR							
55		(ix)	MER			+2	, 0									
		,			ME/S	CHT.ñ	CCET	· CF								
					GE:											
			. –													

18

60

#### (ix) MERKMALE:

DIRPOCIO- OF 4420785415

(A) NAME/SCHLÜSSEL: 3'UTR

(B) LAGE: 790..996

(B) LAGE: 790996		5
(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 7:	<b>60</b>	,
TGGTATAATG ATACAGGAAG TGAAAATCCG AAAGGTTCAG ACGATGAAAA GAGTTTGAGA	60	
CGCATCAATG ATCAGCTTTG AGCTATATGT AAGTCTATTA ATTGATTACT AATAGCAATT	120	
TATGGTATCC TCTGTTCTGC ATATCGACGG TTCTCACGTG ATGATCAGCT TGAGGCTTCG	180 240	10
CGGATAAAGT TCCATCGATT ACTATAAAAC CATCACATTA AACGTTCACT ATAGGCATAC	294	
ACACAGACTA AGTTCAAGTT AGCAGTGACA ATG ATT AAG GGA TTA GGC GAA GTT  Met Ile Lys Gly Leu Gly Glu Val	27.	
Mec Tie Lys Giy Hed Giy Gid Var		
GAT CAA ACC TAC GAT GCG AGC TCT GTC GAG GTT GGC ATT GTC CAC GCG	342	15
Asp Gln Thr Tyr Asp Ala Ser Ser Val Glu Val Gly Ile Val His Ala		
10 15 20 AGA TGG AAC AAG ACT GTC ATT GAC GCT CTC GAC CAA GGT GCA ATT GAG	390	
Arg Trp Asn Lys Thr Val Ile Asp Ala Leu Asp Gln Gly Ala Ile Glu		20
25 30 35 40		
AAA CTG CTT GCT ATG GGA GTG AAG GAG AAG AAT ATC ACT GTA AGC ACC	438	
Lys Leu Leu Ala Met Gly Val Lys Glu Lys Asn Ile Thr Val Ser Thr		25
45 50 55		25
GTT CCA GGT GCG TTT GAA CTA CCA TTT GGC ACT CAG CGG TTT GCC GAG	486	
Val Pro Gly Ala Phe Glu Leu Pro Phe Gly Thr Gln Arg Phe Ala Glu		
60 65 70		30
CTG ACC AAG GCA AGT GGC AAG CAT TTG GAC GTG GTC ATC CCA ATT GGA	534	-
Leu Thr Lys Ala Ser Gly Lys His Leu Asp Val Val Ile Pro Ile Gly		
75 80 85		
GTC CTG ATC AAA GGC GAC TCA ATG CAC TTT GAA TAT ATA TCA GAC TCT	582	35
Val Leu Ile Lys Gly Asp Ser Met His Phe Glu Tyr Ile Ser Asp Ser		
90 95 100	630	
GTG ACT CAT GCC TTA ATG AAC CTA CAG AAG AAG ATT CGT CTT CCT GTC	630	
Val Thr His Ala Leu Met Asn Leu Gln Lys Lys Ile Arg Leu Pro Val		40
103	678	
ATT TTT GGT TTG CTA ACG TGT CTA ACA GAG GAA CAA GCG TTG ACA CGT	0,0	
The Phe Gly Leu Leu Thr Cys Leu Thr Glu Glu Gln Ala Leu Thr Arg		
125 130 135 GCA GGC CTC GGT GAA TCT GAA GGC AAG CAC AAC CAC GGT GAA GAC TGG	726	45
Ala Gly Leu Gly Glu Ser Glu Gly Lys His Asn His Gly Glu Asp Trp		
140 145 150		
GGT GCT GCC GTG GAG ATG GCT GTA AAG TTT GGC CCA CGC GCC GAA	774	50
Gly Ala Ala Val Glu Met Ala Val Lys Phe Gly Pro Arg Ala Glu		
155 160 165		
CAA ATG AAG AAG TGAATATTAA AAAATCACTA CTTAAAATTA ACGTTTTTAT	826	
Gln Met Lys Lys		55
170		
TATGTCTATA TCAAATTCTT ACGTGATAAC TTTTGATTTC GCTTCCTGGA TTGGCGCAAG	886	
GCCTCCCTGT GTCGCAGTTT TTGTTCACGG GTCCACACAG CTCTGTTTTC CCAGAACATA	946	
TCCTCCCAGC CGGCGAACCG GTTAGACGCT TCTGCTGGCG TTCTTATTTT	996	60

(2) INFORMATION ZU SEQ ID NO: 8: (i) SEQUENZ CHARAKTERISTIKA: (A) LÄNGE: 172 Aminosauren 5 (B) ART: Aminosaure (D) TOPOLOGIE: linear (ii) ART DES MOLEKŪLS: Protein (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 8: Met Ile Lys Gly Leu Gly Glu Val Asp Gln Thr Tyr Asp Ala Ser Ser 10 Val Glu Val Gly Ile Val His Ala Arg Trp Asn Lys Thr Val Ile Asp 20 Ala Leu Asp Gln Gly Ala Ile Glu Lys Leu Leu Ala Met Gly Val Lys 15 40 Glu Lys Asn Ile Thr Val Ser Thr Val Pro Gly Ala Phe Glu Leu Pro 55 Phe Gly Thr Gln Arg Phe Ala Glu Leu Thr Lys Ala Ser Gly Lys His 20 75 Leu Asp Val Val Ile Pro Ile Gly Val Leu Ile Lys Gly Asp Ser Met 85 25 His Phe Glu Tyr Ile Ser Asp Ser Val Thr His Ala Leu Met Asn Leu 105 Gln Lys Lys Ile Arg Leu Pro Val Ile Phe Gly Leu Leu Thr Cys Leu 120 30 Thr Glu Glu Gln Ala Leu Thr Arg Ala Gly Leu Gly Glu Ser Glu Gly 135 140 Lys His Asn His Gly Glu Asp Trp Gly Ala Ala Ala Val Glu Met Ala 150 35 Val Lys Phe Gly Pro Arg Ala Glu Gln Met Lys Lys 165 (2) INFORMATION ZU SEQ ID NO: 9: (i) SEQUENZ CHARAKTERISTIKA: 40 (A) LÄNGE: 1511 Basenpaare (B) ART: Nukleinsäure (C) STRANGFORM: Doppel (D) TOPOLOGIE: linear 45 (ii) ART DES MOLEKŪLS: cDNS zu mRNS (iii) HYPOTHETISCH: NEIN (iii) ANTISENSE: NEIN (vi) URSPRÜNLICHE HERKUNFT: 50 (A) ORGANISMUS: Ashbya gossypii (ix) MERKMALE: (A) NAME/SCHLÜSSEL: 5'UTR (B) LAGE: 1..524 55 (ix) MERKMALE: (A) NAME/SCHLÜSSEL: CDS (B) LAGE: 525..1232

60



#### (ix) MERKMALE:

(A) NAME/SCHLÜSSEL: 3'UTR

( m )	LAGE:	1777	1511
(8)	LAGE:	1233.	

					1233												5
(:	xi)	SEQ	UENZ	BESC	HREI	BUNG	: SE	Q ID	NO:	9:				~~~	20020	60	
TGTAT	TCA	AC C	TGGA	GGAT	A AC	GAAA	TITC	CAT	GGCG	CGG	GCGA	TACC	AA C	CCAC	CACCA	120	
CCAGA	TAT.	AA G	ACCA	ATCC	C GG	CGGG	TGTG	CCA	.GCCG	CCA	TCAG	AGAC	AG C	CCAC	CCCCA	180	
AGGCA	TGT	GA A	GTCA	AAAG	G CG	CCAG	CTCC	TTA	TCCG	CTC	CCGC	ACAA	GC A		AAAA	240	• • •
TATCC	CGA	TG A	GCGC	GCCA	G CA	.CCCA	GACG	CTA	CACC	ACC	ACCA	2000 2000	AC T	יבע תיד יבי	TAAAA	300	
GAGCG	CTT	TC C	AGCT	TCTC	A GG	CAGI	TAGC	TCI	ACGA	CAA	AGGA	ACCA	יבא ער איני יבא ער	SCTC	מממשים	360	
GATAG	ACG	CG A	CTTG	CTCA	A CG	ATGI	TTCT	GIG	ACCA	WAC	CAAG		יאת כ	CCCA	CTAAA	420	
GTATA	ATC.	AG A	TAGT	TAGI	'C GI	ATCI	TCTA	GIT	TTAT	TAG	TCAG		א אא.	CAAC	ACCGC	480	16
CATTT	CCT	TA T	GCAT	GTCI	T AC	GAGT	TTAA	AAA	GCTC		GTAG		. W	יארכי	ATGCA	536	
TAGAT	'GGC	AT A	CCGA	AGCC	T AT	ATC	CCCA	TAG	AAGI	IGA	TAGG	Mat	Pho	Thr	Gly		-
												met 1			Gry		
									<b></b>	~~	mm-0			CAT	מרכ	584	20
ATA G	TG	GAA	CAC	ATT	GGC	ACT	GTT	GCT	GAG	TAC	TIG	Clu	AAC	9 CD	Ala		
Ile V	al	Glu	His	Ile		Thr	Vaı	Ala	GIU		Leu	GIU	ASII	ಗಾಗಿ	20		
5					10					15	3.000	220	Cam	ccc		632	2
AGC G	AG	GCA	GGC	GGC	AAC	CGT	GTG	TCA	GIC	CTT	ATC	AAG	GAI	λla	A 3 a		25
Ser G	lu	Ala	Gly		Asn	GŢĀ	Val	ser		Leu	TIE	Dy S	rap	35	7.44		
				25				000	30	mcc	አመጣ	CCA	TICC		GGT	686	כ
CCG A	ATA	CTG	GCG	GAT	TGC	CAC	ATC	GGT.	3.co	505	Tla	λla	CV5	Asn	Glv		
Pro I	le	Leu		Asp	Cys	Hls	TIE	45	ASD	SEL	116	n. u	50		<b>U</b> -2		30
			40			~~~	mmc		~~~	CAT	»GC	תיזיכי		GTC	GGG	72	В
ATC T	:GC	CTG	ACG	GTG	ACG	GAG	The	MP~	31a	yen	Car	Phe	Tvs	Val	Glv		
Ile C	:ys		Thr	vaı	TUL	GIU	60	1111	MIG	ASP	361	65	_, _		<b>-</b>		
ATC G		55	~		com	ma m		NCG.	CAA	GTC	AGC		TGG	AAA	GCT	77	6 35
Ile A	3CA	CCA	GAA	ACA	GTT	TAT	2-4	mh =	Glu	Wa 1	Cer	Ser	TED	Lvs	Ala		
		Pro	GIU	Thr	Val	75	Arg	TILL	Giu	<b>V</b> CL	80						
GGC I	70		3.000		CER		NCG	ccc	ልጥሮ	TCG		GAC	AGG	CGC	TAC	82	4 40
GGC 1	7	AAG	TIO	AAC	LON	Glu	Ara	Ma	Tle	Ser	Asp	Asp	Arq	Arg	Tyr		40
	er	rys	TIE	Woll	90	GIG	u a			95			•	_	100		
85 GGC G		03.0	ma C	CTC		GGC	CAC	GTC	GAC		GTG	GCC	TCT	ATT	GTA	87	2
Gly G	27.5	ui c	TVC	Val	Gln	Glv	His	Val	ASD	Ser	Val	Ala	Ser	Ile	Val		45
GTA G	2 T Å	nrs	TAT	105	G.1.1.	C+3			110					115			73
TCC A	2 (7)	CAG	CAC		ccc	AAC	יניטים	ATC		TTT	AAG	TTT	AAA	CTG	CGC	92	0
Ser A	A ZA	Clu	ui e	ASD	Glv	λsn	Ser	Ile	Asn	Phe	Lys	Phe	Lys	Leu	Arg		
ser r	ary	GIU	120		Gry	11011		125			•		130				50
GAT (	~ * *	GNG			AAG	TAC	СТА			AAG	GGT	TTT	GTG	GCG	ATC	96	
Asp (	212	Clu	TAC	Glu	Lve	TUT	Val	Val	Glu	Lvs	Glv	Phe	val	Ala	Ile		
ASD (	3 T I I	135	+ <b>Y</b> +	GIU	y 3	-1-	140					145	•				
GAC (	ccm	T23	TCC	رشرة	ልርሜ	ርጥል		AAG	ATG	GAT	CCA			TGT	TTC	101	.6 55
Asp (	CIT	GTG	100	T.em	ጥኮተ	Val	Ser	Lvs	Met	ASD	Pro	Asp	Gly	Cys	Phe		
		vaı	267	Ter	TIIL	155		_, _			160	- 2	-	-			
i	150					100											

60

	m z						)										
	~	C A	IC I	CG. A	TG A	Y	A CA	C AC	G CA	G AC	C GC	T GI	'A GC	C -1	T CC	A CTG	1064
	- 1		le S	er M	et I	le Al	a Hi	s Th	r Gl	n Th	r Al	a Va	1 A1	a Le	u Pr	o Leu	2001
		, ,				17	0				17	5				100	
5	AΑ	G CC	CG G	IC GO	ST GC	C CI	C GT	G AA	C AT	A GA	A AC	G GA	т ст	ממידי	c cc	C 330	1110
	Ly	s Pi	O As	p G	ly Al	a Le	u Va	l As	n Il	e Gl	מים נו	τ λc	בט בי	1 80	- 01	y Lys	1112
					18	5				19		± 40	p va	T AS			
	CT	A GI	'A GA	G A	AG CA	G GT	T GC	A CA	ር ጥል	יים מיים	о С 22	m			19	5 A GGT	
10	Le	u Va	1 G1	u Lv	rs G1	n Va	1 A1:	a C1	5 IA		G AA	T GC	G CA	G CT	G GA	A GGT u Gly	1160
				20	00		- A1	a G1.	u Iy	r re	u As	n Al	a Gl			u Gly	
	GA	G AG	C TC		-	G CN		3 Om/	20	- -				21	0		
	G1	ı Se	r Se	r Dr	to	u ca	- 3	- GT	# CT9	J GA	A AG	G AT	T AT	T GA	A TC	C AAG	1208
16			. 21		O Le	u GI	ı Arç	y va.	r rei	ı Gl	u Ar	g Il	e Il	e Gl	u Se	r Lys	
15	CITI	T 00		_		_		220	) .				22	5			
	Ci	ا در	T AG	C AT	C TC	A AA	r aac	G TG	ATTA!	TTAT	ATC:	TTGG	GTG (	CTGT.	ATAT(	CT	1259
	re	TAL	a se	r Il	e Se	r Ası	l Lys	5									
		23	-				235										
20	TA	rgta'	TGTC	TTA	CGAC	TGT (	AATO	CAGAC	G GC	TGG	CAGC	r gg:	מממ	יר א כ	CCM	CACACC	3m 1346
	TC	TCT	CCCG	CGG	TGAT	CAG (	CTTC	TGTT	T TO	CTC	አውሞን	A COT	מ מים		mage.	ACACC BACACC	T 1319
	TGT	TGT	GCC	AAC	GCAA	ACA 1	GGAG	רידכר	יידי כר	יררכיי		2 024		77.0	TAGO	ACACC PAGACC	C 1379
	TGC	CGT	CAAT	GCA	CGAG	GCG A	ACAG	CTCC	א א מי		TAC	D CON	GIC	JAAC	TCGT	'AGACC 'AGCTT	T 1439
25	GTG	ACC	GAGT	CC				.0100	n ne	بررور	الالالالا	r Cr.	rGTC	AAAC	CGCC	CAGCTT	C 1499
					NT 211	SEO.	TD N	m - 1	^								1511
	(2) INFORMATION ZU SEQ ID NO: 10: (i) SEQUENZ CHARAKTERISTIKA:																
	(i) SEQUENZ CHARAKTERISTIKA:  (A) LĀNGE: 235 Aminosāuren																
30									āure	n							
50					ART:												
					POPOI												
		(11	.) AF	T DE	es Mo	LEKŪ	LS:	Prot	ein								
		(Xi	.) SE	QUEN	IZBES	CHRE	IBUN	G: S	EQ I	D NO	: 10	:					•
35	Met	Phe	Thr	: Gly	' Ile	· Val	Glu	His	Ile	Gly	Thr	Val	Ala	Glu	ጥህተ	Lou	
	-				5	1				10					1 5		
	Glu	Asn	Asp	Ala	Ser	Glu	Ala	Glv	Glv	Asn	Glv	Va 1	Sor	17-3	7	<b>77</b> -	
				20	)			-	25		<u></u>	V 41	261	30	rea	TTE	
40	Lys	Asp	Ala	Ala	Pro	Ile	Leu	Δla		Circ	ui o	T1-	<b>01</b>	30	_		
			35					40	2105	Cys	nrs	TTE			Ser	Ile	
	Ala	Cys	Asn	Glv	Ile	Cve	Len		17-1	m>	<b>61</b>	-1	45				
		50				Cy S	55	1111	vai	THE	GIU		Thr	Ala	Asp	Ser	
45	Phe			Glv	Tle	21-		<b>6</b> 1			_	60					
	65	-1 -		OL,	Ile	ATA	PIO	GIU	unr	Val		Arg	Thr	Glu	Val	Ser	
		ш~~	T	31-	<b>63</b>	70					75					80	
	Jer	TIP	гу	Ата	Gly	Ser	Lys	Ile	Asn	Leu	Glu	Arg	Ala	Ile	Ser	Asp	
50					85					90					0.5		
	ASD	Arg	Arg	Tyr	Gly	${\tt Gly}$	His	Tyr	Val	Gln	Gly	His	Val	Asp	Ser	Va l	
				TOO					105					110			
	Ala	Ser	Ile	Val	Ser	Arg	Glu	His	Asp	Glv	Asn	Ser	Tla	A CD	Dho	T	
65			115			-		120	•				125	watt	LIIG	ny 2	
55	Phe	Lys	Leu	Ara	Asp	Gln	Glu		G111	Tues	m	37- 7	143		_		
		130		-	2		135	~ I ~	324	-y 5	TAT		۸gT	GIU	Lys	Gly	
	Phe		Ala	IJe	Asn	Glv	T-1	ea-	T	m\	••-	140	_				
	145			-20	Asp	1 E V	AGT	ser	ren	Inr	val	Ser	Lys	Met	Asp	Pro	
60						150					155					160	

Asp Gly Cys. Phe Tyr Ile Ser Met Ile Ala His Thr Gln Thr Ala Val	
Ala Leu Pro Leu Lys Pro Asp Gly Ala Leu Val Asn Ile Glu Thr Asp	
	5
Val Asn Gly Lys Leu Val Glu Lys Gln Val Ala Gln Tyr Leu Asn Ala	
195	
Gln Leu Glu Gly Glu Ser Ser Pro Leu Gln Arg Val Leu Glu Arg Ile	10
210 215 220	
Ile Glu Ser Lys Leu Ala Ser Ile Ser Asn Lys	
225 230 235	
(2) INFORMATION ZU SEQ ID NO: 11:	15
(i) SEQUENZ CHARAKTERISTIKA:	
(A) LÂNGE: 1596 Basenpaare	
(B) ART: Nukleinsäure	
(C) STRANGFORM: Doppel	20
(D) TOPOLOGIE: linear	20
(ii) ART DES MOLEKŪLS: CDNS ZU MRNS	
(iii) HYPOTHETISCH: NEIN	
(iii) ANTISENSE: NEIN	25
(VI) URSPRÛNLICHE HERKUNFT:	25
(A) ORGANISMUS: Ashbya gossypii	
(ix) MERKMALE:	
(A) NAME/SCHLÜSSEL: 5'UTR	30
(B) LAGE: 1352	
(ix) MERKMALE:	
(A) NAME/SCHLÜSSEL: CDS	
(B) LAGE: 3531093	35
(ix) MERKMALE:	
(A) NAME/SCHLÜSSEL: 3'UTR	
(B) LAGE: 10941596	
(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 11:	40
ACARGACCO CAGGCGCCAG TCCGAGCTGG AGGAGAACGA GGCGGCGGG TGAGGATTA	
GCCCCCTGCC CATGGACGAT GCGGGTATAC AGACGGCGGG TATACAGACG GCGGGTATAC	
CCGAGAGAGG CACCAGGCCG GCTTCCTCCA GCGATGCAAG GAAGAGAAGG GGACCAGAGG 180	
CONSCRETA COCATOTA CONTRACTA CONTRACTA COCA COCA COCA COCA COCA COCA COCA	45
CGGATGAGAA CGAGTTCTCG ATATTATAGA GGCCCCCGTT TCGAGTGATT GGCGTCAAAA 300	
ACGGCTATCT GCCTTCGTCC GCCCCACCA CCCTCGGGAA CACTGGCAAA CC ATG 355	
Met	
1	50
GCG CTA ATA CCA CTT TCT CAA GAT CTG GCT GAT ATA CTA GCA CCG TAC 403	
Ala Leu Ile Pro Leu Ser Gln Asp Leu Ala Asp Ile Leu Ala Pro Tyr	
5 10 15	
TTA CCG ACA CCA CCG GAC TCA TCC GCA CGC CTG CCG TTT GTC ACG CTG 451	55
Leu Pro Thr Pro Pro Asp Ser Ser Ala Arg Leu Pro Phe Val Thr Leu	
20 25 30	

23

60

	AC	G TA	T GCC	CAC	G TC	CTA	A GAT	GC3	r CG	r ATC	GCG	AAC	G CAZ	AAC	GGT	GAA	400
	Th	1		Glr	ı Ser	Let	ı Asp	Ala	a Arg	; Ile	ala	Lys	Glr	LVS	. G1v	GAA Glu	499
		•	~				4.0					A 0					
5	AG	G AC	GTT	' ATI	TCG	CAT	GAG	GAG	ACC	AAG	ACA	A TTI		רבר ב	י יים	CTA	545
			val	Ile	Ser	His	Glu	Glu	Thr	Lys	Thr	Met	ጥከተ	Hic	The second	Leu	547
	•	•				22					60						
	CGC	TAC	CAT	CAI	' AGC	GGC	ATC	CTG	ATI	' GGC	TCC	CCC	י ארא	ccc	CUTA	65 GCG	
10	Arg	туг	His	His	Ser	Gly	Ile	Leu	Ile	Glv	Ser	Gly	. nb~	* *1-	T	GCG Ala	595
					, ,					75					~ ~		
	GAC	GAC	CCG	GAT	CTC	AAT	TGC	CGG	TGG	ACA	ССФ	GC N	ccc	CZC	80		
	Asp	) Asp	Pro	Asp	Leu	Asn	Cys	Ara	Tro	Thr	Pro	Δla	. GCG	GAC 3	ي ي	GCG	643
15				05					90					0=			
	GAT	' TGC	ACC	GAA	CAG	TCT	TCA	CCA	CGA	CCC	ידיני ע	እ ጥረጎ	mac	95	~~~		
	Asp	Cys	Thr	Glu	Gln	Ser	Ser	Pro	Ara	Pro	TIA	TIO	Tan	GAT	G.L.I.	CGG	691
			100					105	9		+15	TTE		ASD	vaı	Arg	
20	GGC	AGA	TGG	AGA	TAC	CGC	GGG	ጥርር	ΔΔΔ	ልጥአ	CAC	ma m	110				
	Gly	Arg	Trp	Arg	Tyr	Ara	Glv	Ser	Tare	TIA	Clu	TAT	CIG	CAT	AAC	CTT	739
		115			•	3	120		Lly 3	TIE	GIU		ren	HIS	Asn	Leu	
	GGC	AAG	GGG	AAG	GCG	CCC	ΔͲΔ	CTC	CTC	200	000	125	<b></b> -				
25	Gly	Lys	Gly	Lvs	Ala	Pro	Tla	Val	Un 1	Mb-	01	GGT	GAG	CCG	GAG	GTC	787
	130	_	•	_		135		AGI	vai	THE		GIĀ	Glu	Pro	Glu		
	CGC	GAA	CTA	GGC	GTC	AGT	ጥልሮ	CTIC	CAC	CERC	140					145	
	Arg	Glu	Leu	Glv	Val	Ser	TAC	Tou	CAG	CTG	GGT	GTC	GAC	GAG	GGT	GGC	835
30			Leu	2	150	561	TYL	nen	GIN		GIĀ	Val	Asp	Glu	Gly	Gly	
	CGC	TTG	ААТ	ፐርር		GNG	ന്നവ	mmm	<b>~~</b>	155					160		
	Arg	Leu	AAT Asn	רבנים	Glv	Clu	TOU	Dh-	GAG	CGA	CTC	TAT	TCT	GAG	CAC	CAC	883
	-		Asn	165	GLY	GIU	neu	hue	GIU	Arg	Leu	Tyr	Ser	Glu	His	His	
35	CTG	GAA			ΔͲʹϹ	GTC .	C22 .		170					175			
	Leu	Glu	AGT Ser	Va 1	Met :	Ual /	Cli	91	GGC	GCG	GAG	GTG	CTC	AAC	CAG	CTG	931
			Ser '		****	AGT (	GIU (	anc ată	GIY	AIA	Glu	Val	Leu	Asn	Gln	Leu	
	CTG	CTG		CA	ርልጥ	አመጠ /	~m~ .	185					190				•
40	Leu	Leu	CGC (	Pro	yen.	Tlo 1	71G (	JAC .	AGT	CTG (	GTG .	ATC	ACG	ATA	GGA	TCC	979
		195	Arg :		ngų.	-TE /	200 200	rsb	ser	Leu			Thr	Ile	Gly	Ser	
			CTG (	3GC (	י מייד			30000				205					
	AAG Lys	Phe	Leu (	2112	Sor 1	CON C	3GT (	51-1 (	GCG (	GTC '	TCA (	CCA	GCT	GAG (	GAG	GTG	1027
45	Lys 210		,		Jer 1	215	arā /	al A	ALA '	val :	Ser 1	Pro .	Ala	Glu (	Glu '	Val	
		СТА	GAG (	ነልጥ ሰ							220				:	225	
	AAC Asn	Leu (	Glu :	lie t	7-1 7	AAC 3	rgg 1	rgg (	CAC (	GGA A	ACA A	AGT (	GAC A	AGT (	STT (	rtg	1075
	Asn :			.13	230	ASD 1	LD J	rp i	His (	Gly :	Thr S	Ser 2	Asp :	Ser \	/al 1	Leu	
50	TGC (	age (	יפפ ר	יחורי כ	יים איים מיים	03.000				235				2	240		
	TGC (	31v 2	ara t	- TC (	i Australia	AGCC	G'I'I'A	T G	ACTG	GTC 11.7	A CTA	GTT	AAAA	CTAT	PTTA	CTC	1130
	-, -,	<u>,</u> .			ııd												
		\		45													
55	CIMIN	CATA	AT TO	CGTC	ACAI	AGC	GTTT	ATC	CCCC	TCGC	CA A	CCG	CTC	T GO	CGT:	rggaa	1190
		-96	30 00	GGGC	ACCI	. CAA	GCGC	TCC	GCAT	የሮርልረ	מיחיר	TOTAL PARTY.	N A COCOC				1250
		u.c	یں دی	TWYC	بالخالخال	AGA	.GGTC	TCT	GACT	لمكتلملي	ת מים	יתו ג גיי					
				WCC.T	ATAA	LTGC	AACT	TGG	ልጥሮር	` ል ጥ/// /	ייתירי מי	VCC 2.0	nmam:				1370
60	1 1 CC	LTTAL	CT	GTAT	CTCT	, TCY	ACAA	CTC	CTTC	TTTT	CT I	CGTC	GCTC	A GI	TTG	TATG	1430

	1400	
TTTTGGCACA AGCTCATGGT GCGTGATATT TACCACCAAA GCTGTTTCGT TGAAAGTCTC	1490	
AATTGTAGCA GGAGCGACGG AGGGAAGCAG TTTCAACGCG CTGGGCGTTA TGCCGTTCTG	1550	
ATATATGAAA ATACCCGTCT GGAAGTTCTT CTCGCCAATG TGGATC	1596	
(2) INFORMATION ZU SEQ ID NO: 12:	5	
(i) SEQUENZ CHARAKTERISTIKA:		
(A) LÂNGE: 246 Aminosäuren		
(B) ART: Aminosäure		
(D) TOPOLOGIE: linear	10	
(ii) ART DES MOLEKÜLS: Protein		
(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 12:		
Met Ala Leu Ile Pro Leu Ser Gln Asp Leu Ala Asp Ile Leu Ala Pro	15	
1 . 5 10 15	13	
Tyr Leu Pro Thr Pro Pro Asp Ser Ser Ala Arg Leu Pro Phe Val Thr		
20 25 30		
Leu Thr Tyr Ala Gln Ser Leu Asp Ala Arg Ile Ala Lys Gln Lys Gly	20	
25 40 45	20	
Glu Arg Thr Val Ile Ser His Glu Glu Thr Lys Thr Met Thr His Tyr		
55 <sup>50</sup>		
Leu Arg Tvr His His Ser Gly Ile Leu Ile Gly Ser Gly Thr Ala Leu	25	
45 70 75	23	
Ala Asp Asp Pro Asp Leu Asn Cys Arg Trp Thr Pro Ala Ala Asp Gly		
85 90 95		
Ala Asp Cys Thr Glu Gln Ser Ser Pro Arg Pro Ile Ile Leu Asp Val	30	
100 105 110		
Arg Gly Arg Trp Arg Tyr Arg Gly Ser Lys Ile Glu Tyr Leu His Asn		
115 120 125		
Leu Gly Lys Gly Lys Ala Pro Ile Val Val Thr Gly Gly Glu Pro Glu	35	j
130 135 140		
Val Arg Glu Leu Gly Val Ser Tyr Leu Gln Leu Gly Val Asp Glu Gly		
145 150 155 160		
Gly Arg Leu Asn Trp Gly Glu Leu Phe Glu Arg Leu Tyr Ser Glu His	40	)
165 170 175		
His Leu Glu Ser Val Met Val Glu Gly Gly Ala Glu Val Leu Asn Gln		
180 185 190		
Leu Leu Leu Arg Pro Asp Ile Val Asp Ser Leu Val Ile Thr Ile Gly	45	5
19S 200 20S		
Ser Lys Phe Leu Gly Ser Leu Gly Val Ala Val Ser Pro Ala Glu Glu		
210 215 220		
Val Asn Leu Glu His Val Asn Trp Trp His Gly Thr Ser Asp Ser Val	50	0
225 230 235 240		
Leu Cys Gly Arg Leu Ala		
245		
	5	5

#### Patentansprüche

1. DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 2 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO: 2, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren.

2. DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 4 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO: 4, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren.

65

onne die enzymatische wirkung des rolypeptids weschillen zu reduzieren.

3. DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 6 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO: 6, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind,

ohne die enzymatische W. Lang des Polypeptids wesentlich zu reduziere.

4. DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 8 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO: 8, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren.

5. DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 10 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO: 10, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind,

ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren.

6. DNA-Sequenzen, die für ein Polypeptid mit der in SEQ ID NO: 12 dargestellten Aminosäuresequenz codieren oder für ein Analoges oder Derivat des Polypeptids gemäß SEQ ID NO: 12, worin eine oder mehrere Aminosäuren deletiert, hinzugefügt oder durch andere Aminosäuren substituiert worden sind, ohne die enzymatische Wirkung des Polypeptids wesentlich zu reduzieren.

7. Expressionsvektor, enthaltend eine oder mehrere DNA-Sequenzen gemäß Anspruch 1 bis 6.

8. Wirtsorganismus der mit einem Expressionssystem gemäß Anspruch 7 transformiert worden ist.

9. Rekombinantes Herstellverfahren für Riboflavin, dadurch gekennzeichnet, daß ein Wirtsorganismus gemäß Anspruch 8 verwendet wird.

Hierzu 7 Seite(n) Zeichnungen

26

5

10

15

20

25

30

35

40

45

50

55

60

- Leerseite -



DE 44 20 785 A1 C 12 N 15/80 5. Oktober 1995

Fig. 1

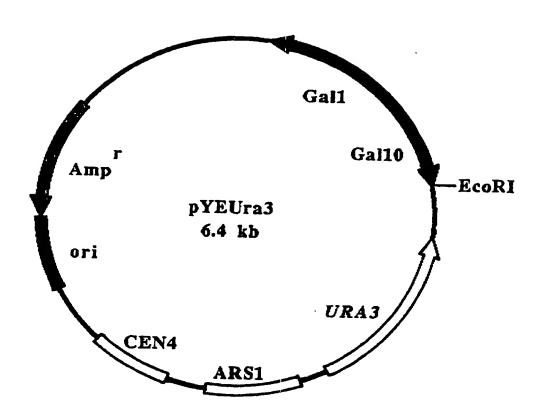


Fig. 2

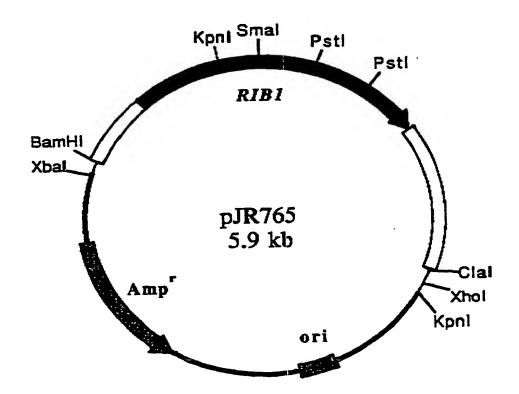
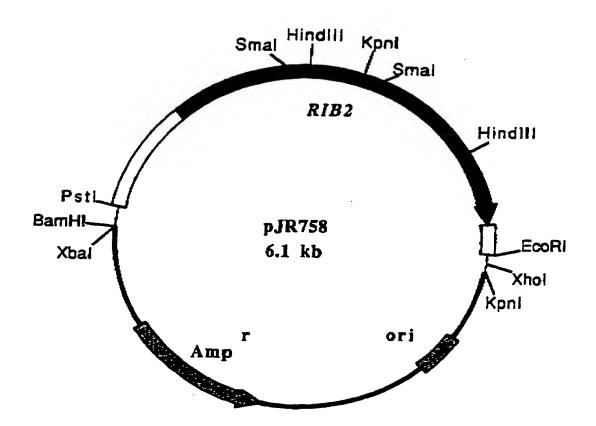


Fig. 3





DE 44 20 785 A1 C 12 N 15/80 5. Oktober 1995

Fig. 4

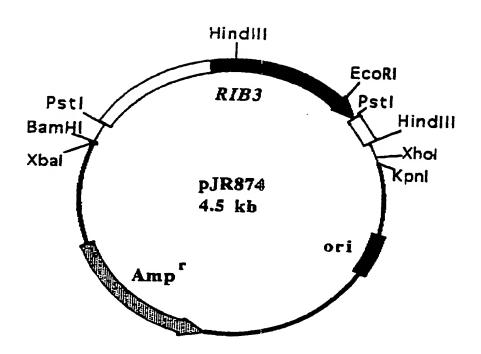


Fig. 5

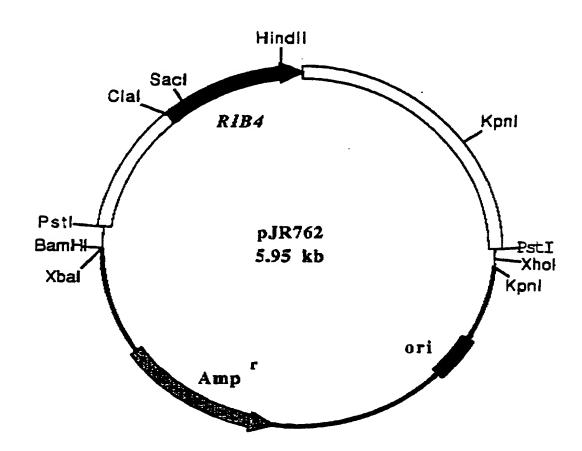


Fig. 6

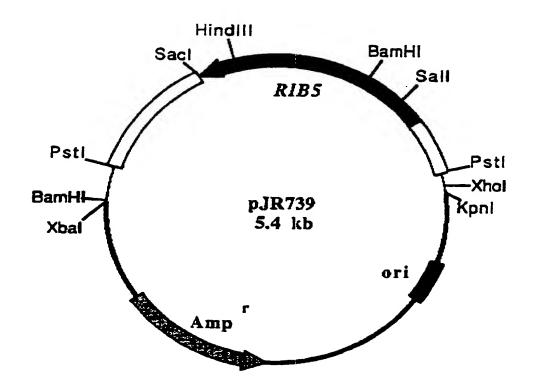


Fig. 7

